QED_3 theory of pairing pseudogap in cuprates: From *d*-wave superconductor to antiferromagnet via an algebraic Fermi liquid

M. Franz
Department of Physics and Astronomy, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z1

Z. Tešanović and O. Vafek
Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218
(Received 24 March 2002; published 29 August 2002)

High-T_c cuprates differ from conventional superconductors in three crucial aspects: the superconducting state descends from a strongly correlated Mott-Hubbard insulator (as opposed to a Fermi liquid), the order parameter exhibits *d*-wave symmetry, and fluctuations play an all important role. We formulate an effective theory of underdoped cuprates within the pseudogap state by taking advantage of these unusual features. In particular, we introduce a concept of ''pairing protectorate'' and we seek to describe various phases within this protectorate by phase disordering a *d*-wave superconductor. The elementary excitations of the protectorate are the Bogoliubov–de Gennes quasiparticles and topological defects in the phase of the pairing field—vortices and antivortices—which appear as quantum and thermal fluctuations. The effective low-energy theory of these elementary excitations is shown to be, apart from intrinsic anisotropy, equivalent to the quantum electrodynamics in $(2+1)$ spacetime dimensions (QED_3). A detailed derivation of this QED_3 theory is given and some of its main physical consequences are inferred for the pseudogap state. As the superconducting order is destroyed by underdoping two possible outcomes emerge: (i) the system can go into a symmetric normal state characterized as an ''algebraic Fermi liquid'' (AFL) before developing antiferromagnetic (AF) order or (ii) a direct transition into the insulating AF state can occur. In both cases the AF order arises spontaneously through an intrinsic ''chiral'' instability of QED_3/AFL. Here we focus on the properties of the AFL and propose that inside the pairing protectorate it assumes the role reminiscent of that played by the Fermi liquid theory in conventional metals. We construct a gauge-invariant electron propagator of the AFL and show that within the $1/N$ expansion it has a non-Fermi-liquid, Luttinger-like form with positive anomalous dimension $\eta' = 16/3\pi^2 N$, where N denotes the number of pairs of nodes. We investigate the effects of Dirac anisotropy by perturbative renormalization group analysis and find that the theory flows into an isotropic fixed point. We therefore conclude that, at long length scales, the AFL is stable against anisotropy.

DOI: 10.1103/PhysRevB.66.054535 PACS number(s): 74.60.Ec, 74.72.−h

I. INTRODUCTION

In the classic tradition of condensed matter physics, the phenomenon of superconductivity is usually described as an instability of a normal metal towards a quantum state in which electrons bind into Cooper pairs.[1] This traditional picture, further fortified within the Eliashberg formalism,[2] has met with spectacular success when applied to conventional, low T_c superconductors. This success is no accident: the BCS-Eliashberg theory uses the Landau theory of Fermi liquids as the underlying description of a normal metal. In turn, the Landau theory of Fermi liquids is one of the most successful theories in physics—by starting from the free gas of fermions it exploits the constraint on the phase space imposed by the Pauli exclusion principle to provide a comprehensive understanding of the low-energy behavior of an *interacting* system. Among its many pleasing features, the Landau theory allows for methodical understanding of the seeds of its own destruction—while the Fermi liquid is hardly the true ground state of any system,[3] we routinely seek to understand such ground states by investigating various instabilities of the Fermi liquid (FL) description, either in the particle-hole or particle-particle channel, few notable examples being spin- and charge-density waves, itinerant ferromagnetism, and, of course, superconductivity. This remarkable success story of the BCS-Eliashberg-FL theory gave birth to the traditional paradigm: ''one must understand the normal state before one can understand the superconductor.''[4,5]

The discovery of high-T_c superconductors (HTS's) and the subsequent efforts to decipher their mysteries altered the above state of affairs dramatically. As experimental information began to pour in, it became rapidly clear that HTS's depart qualitatively from the BCS-Eliashberg-FL orthodoxy in at least three important ways. First, the high-T_c superconducting copper oxides are strongly interacting systems in which correlations play an essential role. Their parent compounds, believed to contain one hole per copper within CuO_2 layers, are far from good metals that simple band-structure theory would predict and are instead insulating Néel antiferromagnets. Even when doped and the long-range antiferromagnetic order had subsided, these materials do not seem to follow the dictates of the Fermi liquid theory, except perhaps in the heavily overdoped regime. Rather, the ''normal'' metal state of the cuprates is anything but, and has been routinely dubbed ''anomalous'' or ''strange.'' In line with the traditional paradigm, a major theoretical effort, mainly inspired by Anderson,[4] has been directed at first understanding this anomalous ''normal'' state as arising from the Hubbard-Mott-Néel antiferromagnetic insulator at half-filling once a small density of mobile dopant holes has been induced in the CuO_2 layers. The physics of such a ''doped Mott insulator'' and particularly the microscopic mechanism through which the high-temperature superconductivity itself is generated from within such a ''normal'' state, even after years of concentrated theoretical and experimental onslaught, remain a deep challenge to the condensed matter physics community.

The second departure from the BCS-Eliashberg-FL orthodoxy[6] has by now been firmly established experimentally: the superconducting state in cuprates is itself of unconventional symmetry.[7] Instead of a usual ''s-wave'' gap function whose magnitude generally varies over the Fermi surface but whose sign does not, the cuprates are ''d-wave'' superconductors. It is believed at the present time that this d-wave symmetry reflects the presence of correlations caused by strong on-site repulsion and is thus a close dynamical relative of the antiferromagnetic state at half-filling. The majority of authors have focused on such a purely electronic ''mechanism'' of superconductivity arising from strong correlations in sharp distinction to the phonon-mediated pairing of the traditional BCS-Eliashberg approach. Such unconventional symmetry of the superconducting state and the ensuing presence of low-energy fermionic excitations near the gap nodes result in a rich phenomenology of cuprates. Most remarkably, however, this phenomenology, at low energies and temperatures and deep inside the superconducting state, seems to fit within the theoretical mold of a model BCS-like d-wave superconductor, quite separately from the detailed microscopic mechanism that is at its origin. We consider this an experimental fact of crucial significance which for the rest of this paper we intend to fully exploit to our advantage.

Finally, more than the intrinsic gap symmetry distinguishes HTS's from their conventional kin. They are also strongly fluctuating systems.[8] While strong fluctuations away from a simple mean-field description are a familiar occurrence in magnets or liquid crystals they are a relatively novel phenomenon in superconductors. In most conventional superconductors, well described by the BCS-Eliashberg-FL theory—the fluctuations are simply not an issue: the dimensionless parameter which controls the deviations from the mean-field theory—the inverse of the product of the BCS coherence length ξ_0 and the Fermi wave vector k_F—is typically as small as 10^{-3} or 10^{-4} and the fluctuation effects are rarely observable.[9] In contrast, early experiments on HTS's clearly established strong fluctuations in numerous physical quantities and various estimates put $(1/\xi_0 k_F)\sim 10^{-1}$ or larger. The strongly fluctuating nature of the superconducting state is particularly pronounced in underdoped cuprates and is rather vividly manifested in recent experiments on the Nernst effect by the Princeton group.[10] The Nernst effect is routinely used to detect superconducting vortex fluctuations and a strong signal observed up to $T_{\rm Nernst}\gg T_c$, with $T_{\rm Nernst}$ comparable to the pseudogap temperature T^*, is most naturally interpreted in those terms. Other experiments have also provided both direct[11] and indirect[12] support for pairing fluctuations far above T_c. The experimental evidence for such fluctuations has been compiled in Ref. 13.

Recently, a different path toward the theory of HTS's was proposed in Refs. 14 and 15. A key aspect of this ''inverted'' approach is the observation that, when compared to its neighbors in the phase diagram of cuprates, the superconducting state appears to be the ''least correlated'' and its fermionic excitations best defined.[15] A tantalizing theoretical feature of this approach is that it holds promise to turn the above triple predicament of strong correlations, fluctuations, and unconventional symmetry in HTS's into an advantage by using the superconducting state as its departure point. The first step in constructing a theory based on this ''inverted'' philosophy is to assume that the most important effect of strong correlations at the basic microscopic level is to build a large pseudogap of d-wave symmetry, which is predominantly pairing in origin, i.e., arises form the particle-particle (p-p) channel. In this regard, we are following the wisdom of the FL theory but are replacing the free electron gas starting point with the free Bogoliubov–de Gennes (BdG) d-wave quasiparticles. Under the umbrella of this large d-wave pseudogap, the low-energy fermionic excitations enjoy a remarkably sheltered existence. Within this ''pairing protectorate,'' they are completely impervious to weak residual short-range interactions left over after the effect of the pseudogap had been built in.

There is an important new element in this parallel with the FL. Our reference state being a superconductor, we must also consider interactions of BdG quasiparticles with relevant collective modes of the pairing pseudogap, i.e., fluctuating thermal and quantum vortex-antivortex pairs. Within the theory of Ref. 15 this interaction is represented by two U(1) gauge fields v_μ and a_μ. Here v_μ describes Doppler shift in quasiparticle energies[16] and has been studied in Refs. 17 and 18. Its effect on low-energy fermions is rather modest since v_μ gains mass from fermions both in a superconductor and in a phase-incoherent pseudogap state. In contrast, the Berry gauge field a_μ, minimally coupled to fermions and encoding the topological frustration inflicted upon BdG quasiparticles by fluctuating vortices, is massless in the pseudogap state and is the main source of strong scattering at low energies.[15] The effective theory was found to take the form equivalent to (2+1)-dimensional (anisotropic) quantum electrodynamics (QED_3).[15] In its symmetric phase, QED_3 is governed by the interacting critical point leading to a non-Fermi-liquid behavior for its fermionic excitations. This ''algebraic'' Fermi liquid[15] (AFL) displaces conventional FL as the underlying theory of the pseudogap state.

The AFL (symmetric QED_3) suffers an intrinsic instability when vortex-antivortex fluctuations and residual interactions become *too strong*. The topological frustration is relieved by the spontaneous generation of mass for fermions, while the Berry gauge field remains massless. In the field theory literature on QED_3 this instability is known as dynamical chiral symmetry breaking (CSB) and is a well-studied and established phenomenon,[19] although few clouds of uncertainty still hover over its more quantitative aspects.[20] In cuprates, the region of such strong vortex fluctuations corresponds to heavily underdoped samples and CSB leads to the spontaneous creation of a whole plethora of nearly degenerate ordered and gapped states from within the AFL.[22] An important check on the internal consistency of the ''inverted'' approach is that the manifold of CSB states contains an incommensurate antiferromagnetic insulator [spin-density

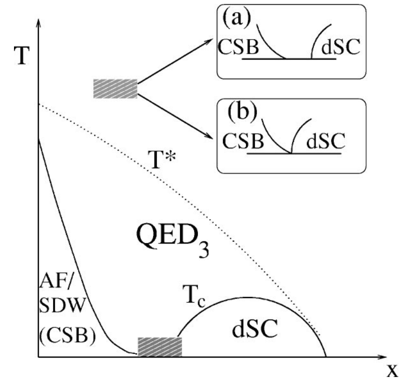

FIG. 1. A schematic representation of the phase diagram of a cuprate superconductor. Below T^* ($\sim T_{\mathrm{Nernst}}$), the symmetric QED_3/AFL replaces the Fermi liquid as the effective low-energy theory. CSB denotes chiral-symmetry-broken states, the most prominent among which is an incommensurate antiferromagnet (SDW). Panel (a) represents the path between dSC and an antiferromagnet via the intervening algebraic Fermi liquid (AFL) ground state. Panel (b) shows a direct dSC-CSB transition with AFL describing $T \ll T^*$ behavior.

wave (SDW)].[21,22] Remarkably, both the "algebraic" Fermi liquid and the SDW and other CSB insulating states arise from one and the same QED_3 theory,[15] echoing the satisfying features of Fermi liquid theory in conventional metals. It therefore appears that the "inverted" approach can be used to advance along the doping axis of the HTS phase diagram (Fig. 1) in the "opposite" direction, from a *d*-wave superconductor all the way to an antiferromagnetic insulator at very low doping, the low-energy physics of the pairing protectorate held under overall control of the symmetric QED_3 (AFL).

In this paper we first present a detailed derivation of the QED_3 theory of the pairing pseudogap state in underdoped cuprates previously introduced in Ref. 15 and then embark on a systematic exploration of its fermionic excitation spectrum and other related properties. To keep the paper at a manageable length we confine ourselves to the the chirally *symmetric* phase of QED_3, i.e., to the AFL and its main properties. The chiral-symmetry-broken phase is discussed separately, in part II of this paper. Section II and Appendix A contain a step-by-step manual on vortex-quasiparticle interactions and how the low-energy physics of such interactions can be given its mathematical formulation in the language of the QED_3 effective theory. In Secs. III and IV we focus on the AFL, give a detailed accounting of vortex-quasiparticle interactions within QED_3, and compute quasiparticle spectral properties within the pseudogap symmetric phase. In Sec. V we then discuss the effects of Dirac anisotropy on the QED_3 infrared fixed point and demonstrate that, to leading order in a $1/N$ expansion, the anisotropic QED_3 scales to an isotropic limit. Finally, we present a brief summary of our results and conclusions in Sec. VI.

In part II of the paper, to appear separately, we introduce the concept of chiral symmetry within the *d*-wave pairing pseudogap state.[21,22] We enumerate different physical states within the chiral manifold and discuss in depth various patterns of CSB (Fig. 1) and fermion mass generation within our QED_3 theory. This is followed by applications to the phase diagram of cuprates within the pseudogap state.

II. PRELIMINARIES AND VORTEX-QUASIPARTICLE INTERACTIONS

A. Protectorate of the pairing pseudogap

Our starting point is the assumption that the pseudogap is predominantly particle-particle or pairing (p-p) in origin and that it has a $d_{x^2-y^2}$ symmetry. This assumption is given mathematical expression in the partition function

$$Z=\int \mathcal{D}\Psi^\dagger(\mathbf{r},\tau)\int \mathcal{D}\Psi(\mathbf{r},\tau)\int \mathcal{D}\varphi(\mathbf{r},\tau)\exp[-S],$$

$$S=\int d\tau\int d^2r\{\Psi^\dagger\partial_\tau\Psi+\Psi^\dagger\mathcal{H}\Psi+(1/g)\Delta^*\Delta\},\qquad(1)$$

where τ is the imaginary time, $\mathbf{r}=(x,y)$, g is an effective coupling constant in the $d_{x^2-y^2}$ channel, and $\Psi^\dagger=(\overline{\psi_\uparrow},\psi_\downarrow)$ are the standard Grassmann variables. The effective Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H}=\begin{pmatrix}\hat{\mathcal{H}}_e & \hat{\Delta}\\ \hat{\Delta}^* & -\hat{\mathcal{H}}_e^*\end{pmatrix}+\mathcal{H}_{\mathrm{res}},\qquad(2)$$

with

$$\hat{\mathcal{H}}_e=\frac{1}{2m}\left(\hat{\mathbf{p}}-\frac{e}{c}\mathbf{A}\right)^2-\epsilon_F,$$

$\hat{\mathbf{p}}=-i\nabla$ (we take $\hbar=1$), and $\hat{\Delta}$ the *d*-wave pairing operator,[23,24]

$$\hat{\Delta}=\frac{1}{k_F^2}\{\hat{p}_x,\{\hat{p}_y,\Delta\}\}-\frac{i}{4k_F^2}\Delta(\{\partial_x,\partial_y\}\varphi),\qquad(3)$$

where $\Delta(\mathbf{r},\tau)=|\Delta|\exp[i\varphi(\mathbf{r},\tau)]$ is the center-of-mass gap function and $\{a,b\}\equiv(ab+ba)/2$. As discussed in Ref. 24 the second term in Eq. (3) is necessary to preserve the overall gauge invariance. Notice that we have rotated $\hat{\Delta}$ from $d_{x^2-y^2}$ to d_{xy} to simplify the continuum limit. $\int\mathcal{D}\varphi(\mathbf{r},\tau)$ denotes the integral over smooth ("spin-wave") and singular (vortex) phase fluctuations. Amplitude fluctuations of $\hat{\Delta}$ are suppressed at or just below T^* and the amplitude itself is frozen at $2\Delta\sim3.56T^*$ for $T\ll T^*$.

The fermion fields $\psi_\uparrow$ and $\psi_\downarrow$ appearing in Eqs. (1) and (2) do not necessarily refer to the bare electrons. Rather, they represent some *effective* low-energy fermions of the theory, already fully renormalized by high-energy interactions, expected to be strong in cuprates due to Mott-Hubbard correlations near half-filling.[4] The precise structure of such fermionic effective fields follows from a more microscopic theory and is not of our immediate concern here—we are only relying on the *absence* of *true* spin-charge separation which allows us to write the effective pairing term (2) in BCS-like form. The experimental evidence that supports this reasoning, at least within the superconducting state and its immediate vicinity, is rather overwhelming.[10–13] Furthermore, $\mathcal{H}_{\mathrm{res}}$

represents the "residual" interactions, i.e., the part dominated by the effective interactions in the p-h channel. Our main assumption is equivalent to stating that such interactions are in a certain sense a "weak" and less important part of the effective Hamiltonian $\mathcal{H}$ than the large pairing interactions already incorporated through $\hat{\Delta}$. As we progress, this notion of the "weakness" of $\mathcal{H}_{\rm res}$ will be defined more precisely and with it the region of validity of our theory.

The above discussion reveals that our theory is at least partly of phenomenological character and it must rely on a more microscopic description to fully define its basic form specified by Eqs. (1) and (2). While it is still far from established just what such a more microscopic description might be in the case of cuprates, various gauge theories of the t-J and related models[25] could all be used for this purpose at the present time. The main role played by such theories is providing reliable values of parameters that feed into the basic formulation (1) and (2). These include, most importantly, the pairing pseudogap Δ itself, the microscopic values of vortex core energies, the strengths of residual interactions, etc. Once these parameters have been supplied through such an external input, it is our task to solve for the low-energy ($\ll\Delta$), long-distance physics of Eqs. (1) and (2). At present, these needed parameters cannot be computed reliably from a fully microscopic approach. We therefore combine the available theoretical arguments with the experimentally determined phase diagram (Fig. 1) and argue that the pseudogap is indeed pairing in origin and that the transition from the superconductor to the pseudogap state must proceed via thermal and quantum unbinding of vortex-antivortex pairs.

Our "inverted" approach is similar in spirit to the Landau theory of Fermi liquids as applied to conventional metals. In Fermi liquid theory the reference state is that of a noninteracting gas of fermions. As the interactions are turned on adiabatically, the Pauli principle severely restricts the available phase space for scattering and many of the general features of free fermion system are preserved, albeit in a renormalized form.[5] In our case, the reference state is a noninteracting system of Bogoliubov–de Gennes fermions. One can think of the pairing pseudogap Δ as being our "Fermi energy" and the highest-energy scale in the problem. While our theory cannot account for the physics at energies (or temperatures) higher than Δ, we will endeavor to show that the low-energy physics can be computed systematically and is parametrized by a handful of material constants whose values can be extracted either from experiments or from a more microscopic theory.

In this context, the residual interactions among the BdG fermions, both those arising from the p-p channel through weak amplitude fluctuations of Δ and the p-h interactions generated by the effective spin fluctuations within the pseudogap state,[14] can be thought of as our version of "Landau interaction parameters" $\{F^{s,a}_{2l+1}\}$. These interactions are generically short ranged and are even less effective in disturbing the coherence of nodal BdG fermions than their FL counterpart. With all of the Fermi surface gapped apart from the nodal regions, such interactions are irrelevant in the renormalization group (RG) sense and can be absorbed into renormalizations of various relevant parameters (more on this in Sec. III). As the system makes the transition from a phase-coherent superconductor to a phase-incoherent pseudogap state, the nodal BdG fermions interacting only through such residual interactions would hardly be noticed—the nodal BdG liquid is a better Fermi liquid than a conventional metal.

There is, however, another source of interactions among BdG quasiparticles which ruins this state of affairs. Our reference state being a superconductor, we must consider the interactions of BdG fermions with collective modes of Δ relevant near the superconductor-"normal" boundary, i.e., fluctuating vortices. As we demonstrate next, these interactions between BdG quasiparticles and fluctuating vortex-antivortex excitations play the central dynamical role in our theory.

B. Phase fluctuations, vortices, quasiparticles, and topological frustration

The global U(1) gauge invariance mandates that the partition function (1) must be independent of the overall choice of phase for $\hat{\Delta}$. We should therefore aim to eliminate the phase $\varphi(\mathbf{r},\tau)$ from the pairing term (2) in favor of $\partial_\mu\varphi$ terms $[\mu=(x,y,\tau)]$ in the fermionic action. For the regular ("spin-wave") piece of φ this is easily accomplished by absorbing a phase factor $\exp(i\frac{1}{2}\varphi)$ into both spin-up and spin-down fermionic fields. This amounts to "screening" the phase of $\Delta(\mathbf{r},\tau)$ (or an "XY phase," as commonly known) by a "half-phase" field (or "half-XY phase") attached to $\psi_\uparrow$ and $\psi_\downarrow$. However, as discussed in Refs. 23 and 24, when dealing with the singular part of φ, such a transformation "screens" physical singly quantized $hc/2e$ superconducting vortices with "half-vortices" in the fermionic fields. Consequently, this "half-angle" gauge transformation must be accompanied by branch cuts in the fermionic fields which originate and terminate at vortex positions and across which the quasiparticle wave function must switch its sign. These branch cuts are a mathematical manifestation of a fundamental physical effect: in the presence of vortices, which are topological defects in the phase of the Cooper pair field and thus naturally bind the elementary flux of $hc/2e$, the motion of quasiparticles is *topologically frustrated*, since their natural elementary flux is twice as large (hc/e). The physics of this topological frustration is at the origin of all nontrivial dynamics discussed in this paper.

Dealing with branch cuts in a fluctuating vortex problem is a rather cumbersome affair due to their nonlocal character and defeats the original purpose of reducing the problem to that of fermions interacting with a *local* fluctuating superflow field, i.e., with $\partial_\mu\varphi$. Instead, in order to avoid the branch cuts, nonlocality, and non-single-valued wave functions, we employ the singular gauge transformation devised in Ref. 23, hereafter referred to as the "FT" transformation:

$$\bar{\psi}_\uparrow\to\exp(-i\varphi_A)\bar{\psi}_\uparrow\,,\qquad \bar{\psi}_\downarrow\to\exp(-i\varphi_B)\bar{\psi}_\downarrow\,, \tag{4}$$

where $\varphi_A+\varphi_B=\varphi$. Here $\varphi_{A(B)}$ is the singular part of the phase due to $A(B)$ vortex defects:

$$\nabla\times\nabla\varphi_{A(B)}=2\pi\hat{z}\sum_i q_i\delta(\mathbf{r}-\mathbf{r}_i^{A(B)}(\tau)), \tag{5}$$

with $q_i=\pm1$ denoting the topological charge of the ith vortex and $\mathbf{r}_i^{A(B)}(\tau)$ its position. The labels A and B represent some convenient but otherwise arbitrary division of vortex defects [loops or lines in $\varphi(\mathbf{r},\tau)$] into two sets, although we will soon discuss many virtues of the *symmetrized* transformation (4) which apportions vortex defects *equally* between sets A and B.[23,24] The transformation (4) "screens" the original superconducting phase φ (or "XY phase") with two ordinary "XY phases" φ_A and φ_B attached to fermions. Both φ_A and φ_B are themselves perfectly legitimate phase configurations of $\Delta(\mathbf{r},\tau)$ but simply with fewer vortex defects. The key feature of the transformation (4) is that it accomplishes "screening" of the physical $hc/2e$ vortices by using only "whole" (i.e., not "halved") vortices in fermionic fields and thus guarantees that the quasiparticle wave functions remain single valued. The topological frustration still remains, being the genuine physical effect of the branch cuts, but is now incorporated directly into the fermionic part of the action:

$$\mathcal{L}'=\overline{\psi}_\uparrow[\partial_\tau+i(\partial_\tau\varphi_A)]\psi_\uparrow+\overline{\psi}_\downarrow[\partial_\tau+i(\partial_\tau\varphi_B)]\psi_\downarrow+\Psi^\dagger\mathcal{H}'\Psi, \tag{6}$$

where the transformed Hamiltonian $\mathcal{H}'$ is

$$\begin{pmatrix} \frac{1}{2m}(\hat{\pi}+\mathbf{v})^2-\epsilon_F & \hat{D} \\ \hat{D} & -\frac{1}{2m}(\hat{\pi}-\mathbf{v})^2+\epsilon_F \end{pmatrix},$$

with $\hat{D}=(\Delta_0/2k_F^2)(\hat{\pi}_x\hat{\pi}_y+\hat{\pi}_y\hat{\pi}_x)$ and $\hat{\pi}=\hat{\mathbf{p}}+\mathbf{a}$.

The singular gauge transformation (4) generates a "Berry" gauge potential

$$a_\mu=\frac{1}{2}(\partial_\mu\varphi_A-\partial_\mu\varphi_B), \tag{7}$$

which describes half-flux Aharonov-Bohm scattering of quasiparticles on vortices. a_μ couples to BdG fermions *minimally* and mimics the effect of branch cuts in quasiparticle-vortex dynamics.[23,24,26] Closed fermion loops in the Feynman path-integral representation of Eq. (1) acquire the (-1) phase factors due to this half-flux Aharonov-Bohm effect just as they would from a branch cut attached to a vortex defect. The topological frustration is now implemented through the fact that $\partial_\mu\varphi_{A(B)}$, minimally coupled to a loop of BdG fermions in the partition function (1) as it winds around the imaginary time direction, generates a phase equal to the advance in the XY phase $\varphi_{A(B)}(\mathbf{r},\tau)$, which inhabits the same spacetime, along the closed path coinciding with the said fermion loop. Naturally, this advance must be an integer multiple of 2π—the overall factor of $\frac{1}{2}$ in the definition of a_μ, Eq. (7), reduces this further to an integer multiple of π. In the end, the phase factors of such fermion loops, which are gauge-invariant quantities, are all equal to ±1 and are furthermore precisely what we would have obtained had we opted for the "half-angle" gauge transformation and implemented the branch cuts.[24]

The above (±1) prefactors of the BdG fermion loops come on top of general and everpresent U(1) phase factors $\exp(i\delta)$, where the phase δ depends on the spacetime configuration of vortices. These U(1) phase factors are supplied by the "Doppler" gauge field

$$v_\mu=\frac{1}{2}(\partial_\mu\varphi_A+\partial_\mu\varphi_B)\equiv\frac{1}{2}\partial_\mu\varphi, \tag{8}$$

which denotes the classical part of the quasiparticle-vortex interaction. The coupling of v_μ to fermions is the same as that of the usual electromagnetic gauge field A_μ and is therefore *nonminimal*, due to the pairing term in the original Hamiltonian $\mathcal{H}$, Eq. (2). It is this nonminimal interaction with v_μ, which we call the Meissner coupling, that is responsible for the U(1) phase factors $\exp(i\delta)$. These U(1) phase factors are "random," in the sense that they are not topological in nature—their values depend on a detailed distribution of superfluid fields of all vortices and "spin waves" as well as on the internal structure of BdG fermion loops, i.e., what is the sequence of spin-up and spin-down portions along such loops. In this respect, while its minimal coupling to BdG fermions means that within the lattice d-wave superconductor model[24] one is naturally tempted to represent the Berry gauge field a_μ as a *compact* U(1) gauge field, the Doppler gauge field v_μ is decidedly *noncompact*, lattice or no lattice.[27] The reader should be advised, however, that the issue of the "compactness" of a_μ versus the "noncompactness" of v_μ constitutes a moot point: a_μ and v_μ as defined by Eqs. (7) and (8) are *not independent* since the discrete spacetime configurations of vortex defects $\{\mathbf{r}_i(\tau)\}$ serve as *sources* for *both*. We will belabor this important issue in the next subsection.

For now, note that all choices of the sets A and B in transformation (4) are completely *equivalent*—different choices represent different singular gauges[24] and v_μ, and therefore $\exp(i\delta)$, are invariant under such transformations. a_μ, on the other hand, changes but only through the introduction of $(\pm)$ unit Aharonov-Bohm fluxes at locations of those vortex defects involved in the transformation. Consequently, the Z_2 style (±1) phase factors associated with a_μ that multiply the fermion loops remain unchanged. We now symmetrize the partition function with respect to this singular gauge by defining a generalized transformation (4) as the sum over all possible choices of A and B, i.e., over the entire family of singular gauge transformations. This is an Ising sum with 2^{N_l} members, where N_l is the total number of vortex defects in $\varphi(\mathbf{r},\tau)$ and is itself yet another choice of the singular gauge. The many benefits of such a symmetrized gauge will be discussed shortly but we stress here that its ultimate function is calculational convenience. What actually matters for the physics is that the original φ be split into two XY phases so that the vortex defects of every distinct topological class are apportioned *equally* between φ_A and φ_B (4).[23,24]

The physics behind this last requirement can be intuitively appreciated as follows: imagine that we simply *prohibit* singly quantized vortex excitations and replace them in Eq. (6) by *doubly quantized* (hc/e) vortices. In this case, the Berry gauge field attaches a multiple of a *full* Aharonov-Bohm flux to each vortex position and all the topological phase factors in front of fermion loops equal unity—this is equivalent to $a_\mu=0$ and a complete absence of topological frustration. *It is then natural to select the gauge which eliminates a_μ altogether*. This is straightforwardly accomplished by screening each doubly quantized (hc/e) vortex defect with a unit-flux Dirac string in $\psi_\uparrow$ and an equivalent one in $\psi_\downarrow$. The resulting sets of A and B vortices are two identical replicas of each other and $a_\mu=1/2\,(\partial_\mu\varphi_A-\partial_\mu\varphi_B)\equiv 0$. Now, while still staying within the same gauge, we allow doubly quantized vortices to relax into energetically more favorable configurations—they will immediately decay into singly quantized vortices and our sets A and B will end up containing an equal number of singly quantized vortices of each distinct topological class. For example, in the case of two-dimensional (2D) thermal fluctuations $\varphi_{A(B)}$ should each contain a half of the original vortices in φ and a half of antivortices. This is readily achieved by including those (anti)vortex variables whose positions are labeled by $\mathbf{r}_i$ with *i even* into φ_A, while the *odd* ones are absorbed into φ_B. The symmetrization is just a convenient mathematical tool that automatically guarantees this goal.

The above symmetrization leads to the new partition function $Z\to\tilde{Z}$:

$$\tilde{Z}=\int \mathcal{D}\tilde{\Psi}^\dagger\int \mathcal{D}\tilde{\Psi}\int \mathcal{D}v_\mu\int \mathcal{D}a_\mu \times\exp\left[-\int_0^\beta d\tau\int d^2r\tilde{\mathcal{L}}\right], \tag{9}$$

in which the half-flux-to-minus-half-flux (Z_2) symmetry of the singular gauge transformation (4) is now manifest:

$$\tilde{\mathcal{L}}=\tilde{\Psi}^\dagger[(\partial_\tau+ia_\tau)\sigma_0+iv_\tau\sigma_3]\tilde{\Psi}+\tilde{\Psi}^\dagger\mathcal{H}'\tilde{\Psi}+\mathcal{L}_0[v_\mu,a_\mu], \tag{10}$$

where $\mathcal{L}_0$ is the "Jacobian" of the transformation given by

$$\exp\left(-\int_0^\beta d\tau\int d^2r\mathcal{L}_0\right)=2^{-N_l}\sum_{A,B}\int \mathcal{D}\varphi(\mathbf{r},\tau) \times\delta[v_\mu-\tfrac{1}{2}(\partial_\mu\varphi_A+\partial_\mu\varphi_B)] \times\delta[a_\mu-\tfrac{1}{2}(\partial_\mu\varphi_A-\partial_\mu\varphi_B)]. \tag{11}$$

Here σ_μ are the Pauli matrices, $\beta=1/T$, $\mathcal{H}'$ is given in Eq. (6), and, for later convenience, $\mathcal{L}_0$ will also include the energetics of vortex core overlap driven by amplitude fluctuations (see Appendix A) and is thus independent of long-range superflow (and of $\mathbf{A}$). We call the transformed quasiparticles $\tilde{\Psi}^\dagger=(\bar{\tilde{\psi}}_\uparrow,\tilde{\psi}_\downarrow)$ appearing in Eq. (10) "topological fermions" (TF's). TF's are the natural fermionic excitations of the pseudogapped normal state. They are electrically neutral and are related to the original quasiparticles by the inversion of transformation (4).

By recasting the original problem in terms of topological fermions we have accomplished our original goal: the interactions between quasiparticles and vortices are now described solely in terms of two local superflow fields:[23]

$$v_{A\mu}=\partial_\mu\varphi_A,\quad v_{B\mu}=\partial_\mu\varphi_B, \tag{12}$$

which we can think of as superfluid velocities associated with the phase configurations $\varphi_{A(B)}(\mathbf{r},\tau)$ of a (2+1)-dimensional XY model with periodic boundary conditions along the τ axis. Our Doppler and Berry gauge fields v_μ and a_μ are linear combinations of $v_{A\mu}$ and $v_{B\mu}$. Note that a_μ is produced *exclusively* by vortex defects since the "spin-wave" configurations of φ can be fully absorbed into v_μ. All that remains is to perform the sum in Eq. (9) over all the "spin-wave" fluctuations and all the spacetime configurations of vortex defects $\{\mathbf{r}_i(\tau)\}$ of this (2+1)-dimensional XY model.

C. "Coarse-grained" Doppler and Berry U(1) gauge fields (v_μ and a_μ) and their physical significance

Unfortunately, exact integration over the phase $\varphi(\mathbf{r},\tau)$ is prohibitively difficult. To proceed by analytic means we must devise some approximate procedure to integrate over the vortex-antivortex positions $\{\mathbf{r}_i(\tau)\}$ in Eq. (9) which will capture the qualitative features of at least the long-distance, low-energy physics of the original problem. This is where our recasting of the problem in terms of BdG fermions interacting with superflow fields $v_{A\mu}$ and $v_{B\mu}$, Eq. (12), will come in handy. A hint as to how to devise such an approximation comes from examining the role of the Doppler gauge field v_μ in the physics of this problem. To illustrate our reasoning and for simplicity, we consider the finite-T case where we can ignore the τ dependence of $\varphi(\mathbf{r},\tau)$. The results are easily generalized, with appropriate modifications, to include quantum fluctuations.

We start by noting that $\mathbf{V}_s=2\mathbf{v}-(2e/c)\mathbf{A}$ is just the physical superfluid velocity,[28] invariant under both $A\leftrightarrow B$ singular gauge transformations (4) and ordinary electromagnetic U(1) gauge symmetry. The superfluid velocity (Doppler) field, swirling around each (anti)vortex defect, is responsible for a vast majority of phenomena that we associate with vortices: long-range interactions between vortex defects, coupling to an external magnetic field and the Abrikosov lattice of the mixed state, the Kosterlitz-Thouless transition, etc. Remarkably, we will show that its role is essential even for the physics discussed in the present paper, although it now appears as a supporting actor to the Berry gauge field a_μ which ultimately occupies the center stage. To see how this comes about, imagine that $\mathbf{a}$ (a_μ) were absent—then, upon integration over topological fermions in Eq. (10), we obtain the following term in the effective Largangian:

$$M^2\left(\nabla\varphi-\frac{2e}{c}\mathbf{A}\right)^2+(\cdots), \tag{13}$$

where $(\cdots)$ denotes higher-order powers and derivatives of $\nabla\varphi-(2e/c)\mathbf{A}$. In the above we have replaced $\mathbf{v}\to(1/2)\nabla\varphi$ to emphasize that the leading term, with the coefficient M^2

proportional to the bare superfluid density, is just the standard superfluid-velocity-squared term of the continuum XY model—the notation M^2 for the coefficient will become clear in a moment. We can now write $\nabla\varphi=\nabla\varphi_{\rm vortex}+\nabla\varphi_{\rm spin\text{-}wave}$ and reexpress the transverse portion of Eq. (13) in terms of (anti)vortex positions $\{\mathbf{r}_i\}$ to obtain a familiar form

$$\rightarrow\frac{M^2}{2\pi}\sum_{\langle i,j\rangle}\ln|\mathbf{r}_i-\mathbf{r}_j| \tag{14}$$

or, by using $\nabla\cdot\nabla\varphi_{\rm vortex}=0$ and $\nabla\times\nabla\varphi_{\rm vortex}=2\pi\rho(\mathbf{r})$, equivalently as

$$\rightarrow\frac{M^2}{2\pi}\int d^2r\int d^2r'\rho(\mathbf{r})\rho(\mathbf{r}')\ln|\mathbf{r}-\mathbf{r}'|, \tag{15}$$

where $\rho(\mathbf{r})=2\pi\Sigma_i q_i\delta(\mathbf{r}-\mathbf{r}_i)$ is the vortex density.

The Meissner coupling of $\mathbf{v}$ to fermions is very strong—it leads to familiar long-range interactions between vortices which constrain vortex fluctuations to a remarkable degree. To make this statement mathematically explicit we introduce the Fourier transform of the vortex density $\rho(\mathbf{q})=\Sigma_i\exp(i\mathbf{q}\cdot\mathbf{r}_i)$ and observe that its variance satisfies

$$\langle\rho(\mathbf{q})\rho(-\mathbf{q})\rangle\propto q^2/M^2. \tag{16}$$

Vortex defects form an *incompressible* liquid—the long-distance vorticity fluctuations are strongly suppressed. This "incompressibility constraint" is naturally enforced in Eq. (9) by replacing the integral over *discrete* vortex positions $\{\mathbf{r}_i\}$ with the integral over a *continuously* distributed field $\overline{\rho}(\mathbf{r})$ with $\langle\overline{\rho}(\mathbf{r})\rangle=0$. The Kosterlitz-Thouless transition and other vortex phenomenology are still maintained in the nontrivial structure of $\mathcal{L}_0[\overline{\rho}(\mathbf{r})]$. But the long-wavelength form of Eq. (13) now reads

$$M^2\left(2\overline{\mathbf{v}}-\frac{2e}{c}\mathbf{A}\right)^2+(\cdots), \tag{17}$$

and can be interpreted as a *massive* action for a transverse U(1) gauge field $\overline{\mathbf{v}}$. The latter is our *coarse-grained* Doppler gauge field defined by

$$\nabla\times\overline{\mathbf{v}}(\mathbf{r})=\pi\overline{\rho}(\mathbf{r}). \tag{18}$$

We have now gone full circle with the Doppler gauge field $\mathbf{v}$. The coarse-graining procedure has made it into a massive U(1) gauge field $\overline{\mathbf{v}}$ whose influence on TF's disappears in the long-wavelength limit. We can therefore drop it from our low-energy fermiology. Hereafter we shall drop the overbar and use the symbol $\mathbf{v}$ for both the actual and coarse-grained quantities, the meaning being obvious from the context.

The real problem also contains the Berry gauge field $\mathbf{a}$ (a_μ) which now must be restored. However, if we are to take advantage of introducing continuous $\rho(\mathbf{r})$ and eventually dispensing with $\mathbf{v}$, $\mathbf{a}$ (a_μ) cannot remain the same gauge field we started with in Eq. (10) when we embarked on our quest to derive the effective theory. Having replaced the Doppler field $\mathbf{v}$ with the *distributed* quantity (18), we *cannot* simply continue to keep $\mathbf{a}$ (a_μ) specified by half-flux Dirac (Aharonov-Bohm) strings located at *discrete* vortex positions $\{\mathbf{r}_i\}$. Instead, $\mathbf{a}$ (a_μ) must be replaced by some new gauge field which reflects the "coarse graining" that has been applied to $\mathbf{v}$—simply put, the Z_2-valued Berry gauge field (7) of the original problem (10) must be *"dressed"* in such a way so as to *best compensate* for the error introduced by "coarse graining" $\mathbf{v}$. In the language of the RG, we must find such "dressing" of the Berry gauge field, i.e., the form and the bare action for $\mathbf{a}$ (a_μ), which renders any such errors irrelevant for the low-energy physics.

The recipe for such a required "dressing" of the Berry gauge field is straightforward in the FT singular gauge—it takes the form of a noncompact U(1) gauge field with a simple Maxwellian action. To see how this comes about note that if we insist on replacing $\mathbf{v}$ by its "coarse-grained" form (18), the only way to achieve this is to "coarse-grain" *both* $\mathbf{v}_A$ and $\mathbf{v}_B$ in the same manner:

$$\nabla\times\mathbf{v}_A\rightarrow2\pi\rho_A,\quad \nabla\times\mathbf{v}_B\rightarrow2\pi\rho_B, \tag{19}$$

where $\rho_{A,B}(\mathbf{r})$ are now continuously distributed densities of $A(B)$ vortex defects [the reader should contrast this with Eqs. (7) and (8)]. This is because the *elementary* vortex variables of our problem are *not* the sources of $\mathbf{v}$ and $\mathbf{a}$; rather, they are the sources of $\mathbf{v}_A$ and $\mathbf{v}_B$. We cannot separately fluctuate or "coarse-grain" the Doppler and the Berry "halves" of a given vortex defect—they are *permanently confined* into a physical $(hc/2e)$ vortex. On the other hand, we *can* independently fluctuate A and B vortices—this is why it was important to use the singular gauge (4) to rewrite the original problem solely in terms of A and B vortices and associated superflow fields (12).

Following this recipe we can now reassemble the coarse-grained Doppler and Berry gauge fields as

$$\mathbf{v}=\frac{1}{2}(\mathbf{v}_A+\mathbf{v}_B),\quad \mathbf{a}=\frac{1}{2}(\mathbf{v}_A-\mathbf{v}_B), \tag{20}$$

with the straightforward generalization to (2+1)D:

$$v=\frac{1}{2}(v_A+v_B),\quad a=\frac{1}{2}(v_A-v_B). \tag{21}$$

The coupling of $\mathbf{v}_A(v_A)$ and $\mathbf{v}_B(v_B)$ to fermions is a hybrid of Meissner and minimal coupling.[27] They contribute a product of U(1) phase factors, $\exp(i\delta_A)\times\exp(i\delta_B)$, to the BdG fermion loops, with both $\exp(i\delta_A)$ and $\exp(i\delta_B)$ being "random" in the sense of the previous subsection. Upon coarse graining $\mathbf{v}_A(v_A)$ and $\mathbf{v}_B(v_B)$ turn into noncompact U(1) gauge fields and therefore $\mathbf{v}(v)$ and $\mathbf{a}(a)$ must as well.

D. Further remarks on the FT gauge

The above "coarse-grained" theory must have the following symmetry: it has to be invariant under the exchange of spin-up and spin-down labels, $\psi_\uparrow\leftrightarrow\psi_\downarrow$, *without* any changes in $\mathbf{v}$ (v). This symmetry ensures that the (arbitrarily preselected) S_z component of the *spin*, which is the same for TF's and real electrons, decouples from the physical superfluid velocity which naturally must couple only to *charge*. When dealing with discrete vortex defects this symmetry is guaranteed by the singular gauge symmetry defined by the family of transformations (4). However, if we replace $\mathbf{v}_A$ and $\mathbf{v}_B$ by their distributed "coarse-grained" versions, the said symmetry is preserved only in the FT gauge. This is seen by considering the effective Lagrangian expressed in terms of coarse-grained quantities:

$$\mathcal{L}\to\tilde{\Psi}^\dagger\partial_\tau\tilde{\Psi}+\tilde{\Psi}^\dagger\mathcal{H}'\tilde{\Psi}+\mathcal{L}_0[\rho_A,\rho_B], \tag{22}$$

where $\mathcal{H}'$ is given by Eq. (6), $\mathbf{v}_A$, $\mathbf{v}_B$ are connected to ρ_A, ρ_B via Eq. (19), and $\mathcal{L}_0$ is independent of τ.

The problem lurks in $\mathcal{L}_0[\rho_A,\rho_B]$—this is just the entropy functional of fluctuating *free* $A(B)$ vortex-antivortex defects and has the following symmetry: $\rho_{A(B)}\to-\rho_{A(B)}$ with $\rho_{B(A)}$ kept unchanged. This symmetry reflects the fact that the entropic "interactions" do not depend on vorticity. Above the Kosterlitz-Thouless transition we can expand:

$$\mathcal{L}_0\to\frac{K_A}{4}(\nabla\times\mathbf{v}_A)^2+\frac{K_B}{4}(\nabla\times\mathbf{v}_B)^2+(\cdots), \tag{23}$$

where the ellipsis denotes higher-order terms and the coefficients $K_{A(B)}^{-1}\to n_l^{A(B)}$ (see Appendix A for details).

The above-discussed symmetry of Hamiltonian $\mathcal{H}'$, Eq. (6), demands that $\mathbf{v}$ and $\mathbf{a}$, Eqs. (20) and (21), be the natural choice for independent distributed vortex fluctuation gauge fields which should appear in our ultimate effective theory. $\mathcal{L}_0$, however, collides with this symmetry of $\mathcal{H}'$—if we replace $\mathbf{v}_{A(B)}\to\mathbf{v}\pm\mathbf{a}$ in Eq. (23), we realize that $\mathbf{v}$ and $\mathbf{a}$ are *coupled* through $\mathcal{L}_0$ in the general case $K_A\neq K_B$:

$$\mathcal{L}_0\to\frac{K_A+K_B}{4}[(\nabla\times\mathbf{v})^2+(\nabla\times\mathbf{a})^2]+\frac{K_A-K_B}{2}(\nabla\times\mathbf{a})\cdot(\nabla\times\mathbf{v}). \tag{24}$$

Therefore, via its coupling to $\mathbf{a}$, the superfluid velocity $\mathbf{v}$ couples to the "spin" of topological fermions and ultimately to the true *spin* of the real electrons. This is an unacceptable feature for the effective theory and seriously handicaps the general "A-B" gauge, in which the original phase is split into $\varphi\to\varphi_A+\varphi_B$, with $\varphi_{A(B)}$ each containing some *arbitrary* fraction of the original vortex defects. In contrast, the symmetrized transformation (4) which apportions vortex defects equally between φ_A and φ_B (Ref. 23) leads to $K_A=K_B$ and to a decoupling of $\mathbf{v}$ and $\mathbf{a}$ at quadratic order, thus eliminating the problem at its root. Furthermore, even if we start with the general "A-B" gauge, the renormalization of $\mathcal{L}_0$ arising from integration over fermions will ultimately drive $K_A-K_B\to0$ and make the coupling of $\mathbf{v}$ and $\mathbf{a}$ *irrelevant* in the RG sense.[29] This argument is actually quite rigorous in the case of quantum fluctuations where the symmetrized gauge (4) represents a fixed point in the RG analysis (see below).[30] Consequently, it appears that the symmetrized singular transformation (4) employed in Eq. (10) is the *preferred* gauge for the construction of the effective low-energy theory.[30,31] In this respect, while all the singular A-B gauges are created equal some are ultimately more equal than others.

The above discussion provides the rationale behind using the FT transformation in our quest for an effective theory. At low energies, the interactions between quasiparticles and vortices are represented by two U(1) gauge fields v and a, Eqs. (20) and (21). The conversion of a from a Z_2-valued to a noncompact U(1) field with Maxwellian action is effected by the confinement of the Doppler to the Berry half of a singly quantized vortex—in the coarse-graining process the phase factors $\exp(i\delta)$ of the noncompact Doppler part "contaminate" the original (±1) factors supplied by a, Eq. (7). This contamination diminishes as doping $x\to0$ since then $v_F/v_\Delta\to0$ and "vortices" are effectively liberated of their Doppler content. In this limit, the pure Z_2 nature of a is recovered and one enters the realm of the Z_2 gauge theory of Senthil and Fisher.[32] In contrast, in the pairing pseudogap regime of this paper where $v_F/v_\Delta>1$ and singly quantized $(hc/2e)$ vortices appear to be the relevant excitations,[33] we expect the effective theory to take the U(1) form described by v and a, Eqs. (20) and (21).

E. Jacobian $\mathcal{L}_0[v_\mu,a_\mu]$ and its Maxwellian form

Having elucidated the origin of the "coarse-grained" U(1) gauge fields v_μ and a_μ and settled on the symmetrized FT transformation (4) as the natural gauge choice for this problem, there remains one more task to be accomplished before we can conclude this section. We need to derive a precise expression for the long-distance, low-energy form of the "Jacobian" $\mathcal{L}_0[v_\mu,a_\mu]$ which serves as the "bare action" for the gauge fields v_μ and a_μ of our effective theory. As shown below, this form is a noncompact Maxwellian whose stiffness K (or inverse "charge" $1/e^2=K$) stands in intimate relation to the helicity modulus tensor of a dSC and, in the pseudogap regime of strong superconducting fluctuations, can be expressed in terms of a *finite* physical superconducting correlation length $\xi_{\rm sc}$: $K\propto\xi_{\rm sc}^2$ [2D] and $K\propto\xi_{\rm sc}$ [(2+1)D]. As we enter the superconducting phase and $\xi_{\rm sc}\to\infty$, $K\to\infty$ as well (or $e^2\to0$), implying that v_μ and a_μ have become massive. Our derivation, the results of which were originally quoted and used in Ref. 15, can be accomplished with remarkably little algebra and holds for the Ginzburg-Landau model, XY model, or any other representation of superconducting fluctuations. This is no accident—the straightforward relationship between the massless (or massive) character of $\mathcal{L}_0[v_\mu,a_\mu]$ and the superconducting phase disorder (or order) is a consequence of rather general physical and symmetry principles.

To make good on the above claim consider first a simple example of an s-wave superconductor with a large gap Δ extending over all of the Fermi surface. We can also view this as a model for the high-energy BdG quasiparticles in dSC, those far removed from the nodes. The action takes a form similar to Eq. (10):

$$\tilde{\mathcal{L}}=\tilde{\Psi}^\dagger[(\partial_\tau+ia_\tau)\sigma_0+iv_\tau\sigma_3]\tilde{\Psi}+\tilde{\Psi}^\dagger\mathcal{H}_s'\tilde{\Psi}+\mathcal{L}_0[v_\mu,a_\mu], \tag{25}$$

but with $\mathcal{H}_s'$ defined as

$$\begin{pmatrix} \frac{1}{2m}(\hat{\pi}+\mathbf{v})^2-\epsilon_F & \Delta \\ \Delta & -\frac{1}{2m}(\hat{\pi}-\mathbf{v})^2+\epsilon_F \end{pmatrix}, \tag{26}$$

with $\hat{\pi}=\hat{\mathbf{p}}+\mathbf{a}$. Here v_μ in the fermionic action (but not in $\mathcal{L}_0$) goes into $v_\mu-(e/c)A_\mu$ when the electromagnetic field is included. The reader might wish to recall here that we have defined $\mathcal{L}_0[v_\mu,a_\mu]$ in this particular way to clearly separate superflow-mediated interactions among vortices, which include A_μ, from entropic effects and short-range amplitude-driven core-overlap interactions, which do not.

1. Thermal vortex-antivortex fluctuations

We can now reap the benefits of this convenient separation. In the language of BdG fermions the system (25) is a large gap "semiconductor" and the Berry gauge field couples to it minimally through "BdG" vector and scalar potentials $\mathbf{a}$ and a_τ. Such a BdG semiconductor is a poor dielectric diamagnet with respect to a_μ. We proceed to ignore its "diamagnetic susceptibility" and also set $a_\tau=0$ to concentrate on thermal fluctuations. All this means is that the Berry gauge field part of the coupling between quasiparticles and vortices in a large-gap *s*-wave superconductor influences the latter only through weak short-range interactions which are unimportant in the region of strong vortex fluctuations near the Kosterlitz-Thouless transition. We can therefore drop $\mathbf{a}$ from the fermionic part of the action and integrate over it to obtain $\mathcal{L}_0$ in terms of physical vorticity $\rho(\mathbf{r})=(\nabla\times\mathbf{v})/\pi$. Additional integration over the fermions produces the effective free energy functional for vortices:

$$\mathcal{F}[\rho]=M^2\left(2\mathbf{v}-\frac{2e}{c}\mathbf{A}^{\text{ext}}\right)^2+(\cdots)+\mathcal{L}_0[\rho], \tag{27}$$

where we have used our earlier notation and have introduced a small external transverse vector potential $\mathbf{A}^{\text{ext}}$. The ellipsis denotes higher-order contributions to the vortex interactions. As discussed earlier in this section, the familiar long-range interactions between vortices lead directly to the standard Coulomb gas representation of the vortex-antivortex fluctuation problem and Kosterlitz-Thouless transition.

The presence of these long-range interactions implies that the vortex system is incompressible, Eq. (16), and long-distance vortex density fluctuations are suppressed. When studying the coupling of BdG quasiparticles to these fluctuations it therefore suffices to expand the "entropic" part:

$$\mathcal{L}_0[\rho]\cong\frac{1}{2}\pi^2K\delta\rho^2+\ldots=\frac{1}{2}K(\nabla\times\mathbf{v})^2+\cdots. \tag{28}$$

The above expansion is justified above T_c since we know that at $T\gg T_c$ we must match the purely entropic form of a noninteracting particle system, $\mathcal{L}_0\propto\delta\rho^2$ (see Appendix A).

To uncover the physical meaning of the coefficient K we expand the free energy F of the vortex system to second order in $\mathbf{A}^{\text{ext}}$. In the pseudogap state gauge invariance demands that F depend only on $\nabla\times\mathbf{A}^{\text{ext}}$:

$$F[\nabla\times\mathbf{A}^{\text{ext}}]=F[0]+\frac{(2e)^2}{2c^2}\chi\int d^2r(\nabla\times\mathbf{A}^{\text{ext}})^2+\cdots. \tag{29}$$

Note that $((2e)^2/c^2)\chi$ is just the diamagnetic susceptibility in the pseudogap state. χ determines the long-wavelength form of the helicity modulus tensor $\Upsilon_{\mu\nu}(q)$ defined as

$$\Upsilon_{\mu\nu}(q)=\Omega\frac{\delta^2F}{\delta A_\mu^{\text{ext}}(q)\,\delta A_\nu^{\text{ext}}(-q)}\Bigg|_{A_\mu^{\text{ext}}\to 0}. \tag{30}$$

The above is the more general form of $\Upsilon_{\mu\nu}(q)$ applicable to uniaxially symmetric 3D and (2+1)D XY or Ginzburg-Landau (GL) models; in 2D only $\Upsilon_{xy}(\mathbf{q})$ appears. In the long-wavelength limit $\Upsilon_{\mu\nu}(q)$ vanishes as $q^2\chi$:

$$\frac{c^2}{(2e)^2}\Upsilon_{\mu\nu}(\mathbf{q})=\chi\epsilon_{\rho\alpha\mu}\epsilon_{\rho\beta\nu}q_\alpha q_\beta+\cdots, \tag{31}$$

for the isotropic case, while for the anisotropic situation $\chi_\perp\neq\chi_\parallel$,

$$\frac{c^2}{(2e)^2}\Upsilon_{\mu\nu}(\mathbf{q})=(\chi_\parallel-\chi_\perp)\epsilon_{z\alpha\mu}\epsilon_{z\beta\nu}q_\alpha q_\beta+\chi_\perp\epsilon_{\rho\alpha\mu}\epsilon_{\rho\beta\nu}q_\alpha q_\beta. \tag{32}$$

$\epsilon_{\alpha\beta\gamma}$ is the Levi-Civitá symbol, summation over repeated indices is understood, and index z of the anisotropic 3D GL or XY model is replaced by τ for the (2+1)D case. $(2e)^2/c^2$ is factored out for later convenience.

So what is K? Let us compute the helicity modulus of the problem explicitly. This is done by absorbing the small transverse vector potential $\mathbf{A}^{\text{ext}}$ into $\mathbf{v}$ in Eqs. (25) and (26) and integrating over the new variable $\mathbf{v}-(e/c)\mathbf{A}^{\text{ext}}$. The hecility modulus tensor measures the screening properties of the vortex system. In a superconductor, with topological defects bound in vortex-antivortex dipoles, there is no screening at long distances. This translates into a Meissner effect for $\mathbf{A}^{\text{ext}}$. When the dipoles unbind and some free vortex-antivortex excitations appear screening is now possible over all length scales and there is no Meissner effect for $\mathbf{A}^{\text{ext}}$. Information on the presence or absence of such screening is actually stored entirely in $\mathcal{L}_0$, where $\mathbf{A}^{\text{ext}}$ reemerges after the above change of variables. We finally obtain

$$4\chi=K-\frac{K^2\pi^2}{T}\lim_{|\mathbf{q}|\to 0^+}\int d^2re^{i\mathbf{q}\cdot\mathbf{r}}\langle\delta\rho(\mathbf{r})\delta\rho(0)\rangle, \tag{33}$$

where the thermal average $\langle\cdots\rangle$ is over the free energy (27) with $\mathbf{A}^{\text{ext}}=0$ or, equivalently, (25). In the normal phase only the first term contributes in the long-wavelength limit, the second being down by an extra power of q^2 courtesy of long-range vortex interactions. Consequently, $K=4\chi$ [the factor of 4 is due to the fact that the true superfluid velocity is $2\mathbf{v}$ (Ref. 28)].

We see that in the pseudogap phase, with free vortex defects available to screen, $\mathcal{L}_0[\mathbf{v}]$ takes on a massless Maxwellian form, the stiffness of which is given by the diamagnetic susceptibility of a strongly fluctuating superconductor.

In the fluctuation region χ is given by the superconducting correlation length $\xi_{\rm sc}$ (Ref. 34):

$$K=4\chi=4\mathcal{C}_2T\xi_{\rm sc}^2, \tag{34}$$

where $\mathcal{C}_2$ is a numerical constant, intrinsic to a 2D GL, XY, or some other model of superconducting fluctuations.

As we approach T_c, $\xi_{\rm sc}\to\infty$ and the stiffness of the Maxwellian term (28) diverges. This can be interpreted as the Doppler gauge field becoming massive. Indeed, immediately below the Kosterlitz-Thouless transition at T_c, $\mathcal{L}_0\cong m_v^2\mathbf{v}^2+\cdots$, where $m_v\ll M$ and $m_v\to 0$ as $T\to T_c^-$. This is just a reflection of the helicity modulus tensor now becoming *finite* in the long-wavelength limit,

$$\Upsilon_{xy}\to\frac{4e^2}{c^2}\frac{m_v^2M^2}{4M^2+m_v^2}.$$

Topological defects are now bound in vortex-antivortex pairs and cannot screen, resulting in the Meissner effect for $\mathbf{A}^{\rm ext}$. The system is a superconductor and $\mathbf{v}$ had become massive.

Returning to a d-wave superconductor, we can retrace the steps in the above analysis but we must replace $\mathcal{H}_s'$ in Eq. (25) with $\mathcal{H}'$, Eq. (6). Now, instead of a large-gap BdG "semiconductor," we are dealing with a narrow-gap "semiconductor" or BdG "semimetal" because of the low-energy nodal quasiparticles. This means that we must restore the Berry gauge field $\mathbf{a}$ to the fermionic action since the contribution from nodal quasiparticles makes its BdG "diamagnetic susceptibility" very large, $\chi_{\rm BdG}\sim 1/T\gg 1/T^*$ (see the next section). The long-distance fluctuations of $\mathbf{v}$ and $\mathbf{a}$ are now *both* strongly suppressed, the former through incompressibility of the vortex system and the latter through $\chi_{\rm BdG}$. This allows us to expand Eq. (23):

$$\mathcal{L}_0\cong\frac{K_A}{4}(\nabla\times\mathbf{v}_A)^2+\frac{K_B}{4}(\nabla\times\mathbf{v}_B)^2+(\cdots), \tag{35}$$

where $K_A=K_B=K$ is mandated by the FT singular gauge. Since in our gauge the fermion spin and charge channels decouple, $\mathbf{A}^{\rm ext}$ still couples only to $\mathbf{v}$ and the above arguments connecting K to the helicity modulus and diamagnetic susceptibility χ follow through. This finally gives the Maxwellian form of Ref. 15:

$$\mathcal{L}_0\to\frac{K}{2}(\nabla\times\mathbf{v})^2+\frac{K}{2}(\nabla\times\mathbf{a})^2, \tag{36}$$

where K is still given by Eq. (34). Note, however, that $\xi_{\rm sc}$ of a d-wave and of an s-wave superconductor are two rather *different* functions of T, x, and other parameters of the problem, due to strong Berry gauge field renormalizations of vortex interactions in the d-wave case. Nonetheless, as $T\to T_c$, the Kosterlitz-Thouless critical behavior remains unaffected since $\chi_{\rm BdG}$, while large, is still finite at all finite T.

Just as advertised, we have shown that $\mathcal{L}_0[\rho_A,\rho_B]=\mathcal{L}_0[\mathbf{v},\mathbf{a}]$ in the pseudogap state takes on the massless Maxwellian form (36), with the stiffness K set by the true superconducting correlation length $\xi_{\rm sc}$, Eq. (34). This result holds as a general feature of our theory irrespective of whether one employs a Ginzburg-Landau theory, XY model, vortex-antivortex Coulomb plasma, or any other description of strongly fluctuating dSC, as long as such a description properly takes into account vortex-antivortex fluctuations and reproduces Kosterlitz-Thouless phenomenology. In Appendix A we show that within the continuum vortex-antivortex Coulomb plasma model,

$$\mathcal{L}_0\to\frac{T}{2\pi^2n_l}[(\nabla\times\mathbf{v})^2+(\nabla\times\mathbf{a})^2], \tag{37}$$

where n_l is the average density of *free* vortex and antivortex defects. Comparison with Eq. (36) allows us to identify $\xi_{\rm sc}^{-2}\leftrightarrow 4\pi^2\mathcal{C}_2n_l$.[15]

2. Quantum fluctuations of (2+1)D vortex loops

The above results can be generalized to quantum fluctuations of spacetime vortex loops. The superflow fields $v_{A(B)\mu}$ satisfy $(\partial\times v_{A(B)})_\mu=2\pi j_{A(B)\mu}$, where $j_{A(B)\mu}(x)$ are the coarse-grained vorticities associated with $A(B)$ vortex defects and $\langle j_{A(B)\mu}\rangle=0$. The topology of vortex loops dictates that $j_{A(B)\mu}(x)$ be a purely transverse field, i.e., $\partial\cdot j_{A(B)}=0$, reflecting the fact that loops have no starting or ending point. Again, we begin with a large-gap s-wave superconductor and use its poor BdG diamagnetic and dielectric nature to justify dropping the Berry gauge field a_μ from the fermionic part of Eq. (25) and integrating over a_μ in the "entropic" part containing $\mathcal{L}_0[v_\mu,a_\mu]$.

The integration over the fermions contains an important novelty specific to the (2+1)D case: the appearance of Berry phase terms for quantum vortices as they wind around fermions. Such a Berry phase is the consequence of the first-order time derivative in the original fermionic action (1). If we think of spacetime vortex loops as world lines of some relativistic quantum bosons dual to the Cooper pair field $\Delta(\mathbf{r},\tau)$, as we do in Appendix A, then these bosons see Cooper pairs and quasiparticles as sources of "magnetic" flux.[14] At the mean-field level, this translates into a dual "Abrikosov lattice" or a Wigner crystal of holes in a dual superfluid. Accordingly, the nonsuperconducting ground state in the pseudogap regime will likely contain a weak charge modulation—the modulation is made weak by the same strong fluctuations that make $T_c\ll T^*(T_{\rm Nernst})$. The focus of the present paper being a symmetric AFL description of the pseudogap, we postpone the discussion of this point to part II and will ignore it for the rest of this paper. This is justified by the fact that a_μ does not couple to charge directly and is quantitatively valid in the window $T,\Delta_f\ll(\omega,v_{F,\Delta}|\mathbf{q}|)\ll T^*$, where Δ_f is any small gap in the nodal TF spectrum produced by the said charge modulation.

Hereafter, we blissfully turn a blind eye to the above subtleties and assume that the transition from a dSC into a pseudogap phase proceeds via the unbinding of vortex loops of a (2+1)D XY model or its GL counterpart or, equivalently, an anisotropic 3D XY or GL model where the role of imaginary time is taken on by a third spatial axis z.[35] Having learned all we really need from an earlier 2D example we can now integrate the fermions to obtain the effective Lagrangian for coarse-grained spacetime loops $[\pi j_\mu=(\partial\times v)_\mu]$:

$$\mathcal{L}[j_\mu]=M_\mu^2\left(2v_\mu-\frac{2e}{c}A_\mu^{\rm ext}\right)^2+(\cdots)+\mathcal{L}_0[j_\mu], \quad (38)$$

where $M_x=M_y=M$ and $M_\tau=M/c_s$, with $c_s\sim v_F$ being the effective "speed of light" in the vortex loop spacetime. The incompressibility condition reads $\langle j_\mu(q)j_\nu(-q)\rangle\sim q^2[\delta_{\mu\nu}-(q_\mu q_\nu/q^2)]$ and in the pseudogap state permits the expansion

$$\mathcal{L}_0[j_\mu]\cong\frac{1}{2}\pi^2K_\tau j_\tau^2+\frac{1}{2}\pi^2\sum_{i=x,y}K_i j_i^2, \quad (39)$$

where $K_x=K_y=K\neq K_\tau$. Using an analogy with the uniaxially symmetric anisotropic 3D XY (or GL) model we can expand the ground-state energy in the manner of Eq. (29):

$$E[\partial\times A^{\rm ext}]=E[0]+\frac{(2e)^2}{2c^2}\sum_{\perp,\parallel}\chi_{\perp,\parallel}\times\int d^3x(\partial\times A^{\rm ext})^2_{\perp,\parallel}, \quad (40)$$

with $\chi_\perp=\chi_x=\chi_y=\chi$ and $\chi_\parallel=\chi_\tau\neq\chi$. Note that the form of $\mathcal{L}_0$, Eq. (39), follows directly from the requirement that there be infinitely large vortex loops, resulting in vorticity fluctuations over all distances. Combined with Eq. (40) and then translated to the language of a (2+1)D XY (GL) model it tells us something already familiar: upon the transition to the pseudogap state generated by the vortex loop unbinding, the superconductor has turned into an insulator.[14] χ and χ_τ determine the diamagnetic and dielectric susceptibilities of this insulating pseudogap state.

The explicit computation of $\Upsilon_{\mu\nu}(q)$, Eqs. (30) and (32), leads to

$$4\chi_{i,\tau}=K_{i,\tau}-K_{i,\tau}^2\pi^2\lim_{q\to 0^+}\int d^3x e^{iq\cdot x}\langle j(x)_{i,\tau}j(0)_{i,\tau}\rangle, \quad (41)$$

where the second term is again eliminated by the incompressibility of the vortex system. This results in[15]

$$K=4\chi=4\mathcal{C}_3\xi_\tau, \quad K_\tau=4\chi_\tau=4\mathcal{C}_3\frac{\xi_{\rm sc}^2}{\xi_\tau}, \quad (42)$$

where we used the result for the anisotropic 3D XY or GL model: $\chi_\perp=\mathcal{C}_3T\xi_\parallel$, $\chi_\parallel=\mathcal{C}_3T\xi_\perp^2/\xi_\parallel$, $\mathcal{C}_3$ being the intrinsic numerical constant for those models ($\mathcal{C}_3\neq\mathcal{C}_2$). In the case of (2+1)D vortex loops $\xi_\tau\propto\xi_{\rm sc}$ since our adopted model has the dynamical critical exponent $z=1$. Thus, we again encounter a massless Maxwellian form (39) whose stiffness diverges as we approach a superconductor except now this divergence is linear in the superconducting correlation length, $K\propto\xi_{\rm sc}$.

The application to a d-wave superconductor is straightforward: the nodal structure of BdG quasiparticles in dSC helps along the way by providing an anomalous stiffness for the Berry gauge field a_μ—upon integration over the nodal fermions the following term emerges in the effective action:

$$\frac{1}{2}\chi_{\rm BdG}(q)(q^2\delta_{\mu\nu}-q_\mu q_\nu)a_\mu(q)a_\nu(-q)\sim|q|[\delta_{\mu\nu}-(q_\mu q_\nu/q^2)]a_\mu(q)a_\nu(-q). \quad (43)$$

In the terminology of our chimerical BdG "semimetal," the "susceptibility" $\chi_{\rm BdG}$ is not merely very large; it diverges, $\chi_{\rm BdG}\sim 1/q$, as $q\to 0$, and is computed in detail in Sec. IV. We use this pleasing fact to observe that we are fully justified in expanding $\mathcal{L}_0[j_{A\mu},j_{B\mu}]$ and retaining only quadratic terms as long as we keep a safe distance from the actual phase transition:

$$\mathcal{L}_0\cong\frac{K_{A\mu}}{4}(\partial\times v_A)_\mu^2+\frac{K_{B\mu}}{4}(\partial\times v_B)_\mu^2+(\cdots), \quad (44)$$

where $K_{A\mu}=K_{B\mu}=K_\mu$ is again assured by our choice of the FT singular gauge (4) and (11).

The above reasoning merits an amusing aside: beside the ubiquitous incompressibility $(\langle(\partial\times v)_\mu(\partial\times v)_\nu\rangle\sim q^2[\delta_{\mu\nu}-(q_\mu q_\nu/q^2)])$ the integration over nodal fermions now also occasions diverging $\chi_{\rm BdG}$, implying $\langle(\partial\times a)_\mu(\partial\times a)_\nu\rangle\sim q[\delta_{\mu\nu}-(q_\mu q_\nu/q^2)]$. Had we chosen a singular gauge in which the sets A and B were not equivalent, like the Anderson gauge,[29] and therefore $K_{A\mu}\neq K_{B\mu}$, the ensuing $(\partial\times v)\cdot(\partial\times a)$ coupling in $\mathcal{L}_0$, Eqs. (44) and (24), would now be driven to *zero* in the long-distance limit as an *irrelevant* operator in the RG sense.[36] The reason is simple: the coupling of gauge fields v_μ and a_μ to topological fermions mandates that they decouple at the quadratic (harmonic) level due to decoupling of "charge" and "spin" channels for TF. Such a coupling of v_μ and a_μ can only arise from $\mathcal{L}_0$ by our uninformed choice of a singular gauge. Since both $(\partial\times v)^2$ and $(\partial\times a)^2$ terms in Eq. (44) are strongly relevant due to the diverging contributions they receive from fermions, the coupling constant in front of $(\partial\times v)\cdot(\partial\times a)$, proportional to $K_{A\mu}-K_{B\mu}$, is driven to zero under repeated applications of the RG transformation. Therefore, the FT gauge (4) and (11), specifically designed to ensure $K_{A\mu}-K_{B\mu}=0$ at the very start, is recovered as an RG fixed point.[30]

Finally, we rewrite Eq. (44) in terms of $v_\mu,a_\mu=(1/2)\times(v_{A\mu}\pm v_{B\mu})$,

$$\mathcal{L}_0\to\frac{K_\mu}{2}(\partial\times v)_\mu^2+\frac{K_\mu}{2}(\partial\times a)_\mu^2, \quad (45)$$

and observe that the fact that $A_\mu^{\rm ext}$ couples only to v_μ means that the expression (42) for K_μ is still valid. Of course, $\xi_{\rm sc}(x,T)$ is now truly different from its s-wave counterpart, including a possible difference in the critical exponent, since the coupling of d-wave quasiparticles to the Berry gauge field is marginal at the RG engineering level and may change the quantum critical behavior of the superconductor-pseudogap (insulator) transition.

III. QED$_3$: A LOW-ENERGY EFFECTIVE THEORY OF THE PSEUDOGAP STATE

We have now elucidated the nature of the coupling of our two gauge fields v_μ and a_μ to TF's and have specified their

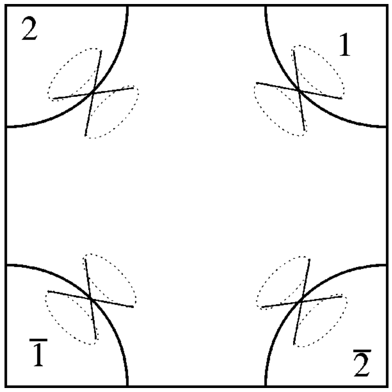

FIG. 2. Schematic representation of the Fermi surface of the cuprate superconductors with the indicated nodal points of the $d_{x^2-y^2}$ gap.

''bare'' thermal and quantum dynamics encoded in $\mathcal{L}_0[v_\mu, a_\mu]$, Eqs. (36) and (45). To make further progress toward our ultimate goal of describing the low-energy fermiology in the pseudogap state, we now focus our attention on the nodal quasiparticle excitations of the Hamiltonian $\mathcal{H}'$, Eq. (6). This will enable us to apply the machinery of the perturbative RG to nodal (massless) TF's and rid our effective theory of all remaining excess baggage.

A. Farewell to v_μ and residual interactions

As indicated in Fig. 2, the low-energy quasiparticles are located at the four nodal points of the d_{xy} gap function: $\mathbf{k}_{1,\bar{1}}=(\pm k_F,0)$ and $\mathbf{k}_{2,\bar{2}}=(0,\pm k_F)$, hereafter denoted as $(1,\bar{1})$ and $(2,\bar{2})$, respectively. To focus on the leading low-energy behavior of the fermionic excitations near the nodes we follow the standard procedure[37] and linearize the Lagrangian (10). To this end we write our TF spinor $\tilde{\Psi}$ as a sum of four nodal Fermi fields,

$$\tilde{\Psi}=e^{i\mathbf{k}_1\cdot\mathbf{r}}\tilde{\Psi}_1+e^{i\mathbf{k}_{\bar{1}}\cdot\mathbf{r}}\sigma_2\tilde{\Psi}_{\bar{1}}+e^{i\mathbf{k}_2\cdot\mathbf{r}}\tilde{\Psi}_2+e^{i\mathbf{k}_{\bar{2}}\cdot\mathbf{r}}\sigma_2\tilde{\Psi}_{\bar{2}}. \tag{46}$$

The σ_2 matrices have been inserted here for convenience: they ensure that we eventually recover the conventional form of the QED_3 Lagrangian. (Without the σ_2 matrices the Dirac velocities at $\overline{1},\overline{2}$ nodes would have been negative.) Inserting $\tilde{\Psi}$ into Eq. (10) and systematically neglecting the irrelevant higher order derivatives,[37] we obtain a nodal Lagrangian of the form

$$\mathcal{L}_D=\sum_{l=1,\bar{1}}\tilde{\Psi}_l^\dagger[D_\tau+iv_FD_x\sigma_3+iv_\Delta D_y\sigma_1]\tilde{\Psi}_l + \sum_{l=2,\bar{2}}\tilde{\Psi}_l^\dagger[D_\tau+iv_FD_y\sigma_3+iv_\Delta D_x\sigma_1]\tilde{\Psi}_l+\mathcal{L}_0[a_\mu], \tag{47}$$

where $D_\mu=\partial_\mu+ia_\mu$ denotes the covariant derivative and $\mathcal{L}_0[a_\mu]$ is given by Eq. (45):

$$\mathcal{L}_0[a_\mu]=\frac{1}{2}K_\mu(\partial\times a)^2_\mu\equiv\frac{1}{2e_\mu^2}(\partial\times a)^2_\mu. \tag{48}$$

$v_F=\partial\epsilon_{\mathbf{k}}/\partial\mathbf{k}$ denotes the Fermi velocity at the node and $v_\Delta=\partial\Delta_{\mathbf{k}}/\partial\mathbf{k}$ denotes the gap velocity. Note that v_F and v_Δ already contain renormalizations coming from high-energy interactions and are effective material parameters of our theory. Similarly, $K_\mu=1/e_\mu^2$, derived in the previous section in terms of $\xi_{sc}(x,T)$, are treated as adjustable parameters which are matched to experimentally available information on the range of superconducting correlations in the pseudogap state.

The Doppler gauge field v_μ has disappeared from the above expression. After informing us on how to properly ''coarse-grain'' the theory and dressing our Z_2-valued Berry gauge field in its ultimate U(1) Maxwellian outfit, the time has come to drop v_μ, its eventual demise caused by the Meissner coupling to BdG fermions discussed in the previous section. After being ''screened'' by high-energy and nodal TF's it is rendered massive both in the superconducting and pseudogap states and unimportant for low-energy physics. Its legacy lives on, however, having given birth to the U(1) noncompact character of a_μ.

In contrast, the Berry gauge field a_μ remains *massless* in the pseudogap state, as it cannot acquire mass by coupling to the fermions. As seen from Eq. (47), a_μ couples minimally to the Dirac fermions and therefore its massless character is protected by gauge invariance. Physically, one can also argue that a_μ couples to the TF *spin* three-current—in a spin-singlet *d*-wave superconductor SU(2) spin symmetry must remain unbroken, thereby ensuring that a_μ remains massless. Its massless Maxwellian dynamics (48) in the pseudogap state can therefore be traced back to the topological state of spacetime vortex loops and directly reflects the absence of true superconducting order (Sec. II) or, equivalently, the presence of a ''vortex loop condensate'' and dual order (Appendix A).

We have also dispensed with the residual interactions represented by $\mathcal{H}_{res}$ in Eq. (2). These interactions are generically short-ranged contributions from the particle-hole (p-h) and amplitude fluctuations part of the p-p channel and in our new notation are exemplified by

$$\mathcal{H}_{res}\to\frac{1}{2}\sum_{l,l'}I_{ll'}\tilde{\Psi}_l^\dagger\tilde{\Psi}_l\tilde{\Psi}_{l'}^\dagger\tilde{\Psi}_{l'}. \tag{49}$$

The effective vertex $I_{ll'}$ has a scaling dimension -1 at the engineering level. This follows from the RG analysis near the massless Dirac points which sets the dimension of $\tilde{\Psi}_l$ to $[\text{length}]^{-1}$. The implication is that $I_{ll'}$ is *irrelevant* for low-energy physics in the perturbative RG sense and we are therefore well within our rights in setting $I_{ll'}\to 0$. However, the residual interactions will not go so quietly into the night: such interactions are known to become important if stronger than some critical value I_c.[38,40] In the present theory, this is bound to happen in the severely underdoped regime and at half-filling, as $x\to 0$.[22] In this case, the residual interactions are becoming large and comparable in scale to the pairing pseudogap Δ, and are likely to cause chiral symmetry breaking which leads to spontaneous mass generation for massless Dirac fermions. The CSB and its variety of patterns in the context of the theory (47) and (48) in underdoped cuprates were discussed in Refs. 21 and 22 and are the subject of part II of this paper. Here, where we have limited ourselves to the chirally symmetric phase, we assume $I_{ll'}<I_c$, which we expect to be the case for moderate underdoping, and set $I_{ll'}\to 0$.

In the end, we are left with Eqs. (47) and (48) as our effective low-energy theory for nodal TF's. This theory, derived previously in Ref. 15, is the chief dynamical muscle behind the physics discussed in this paper. It describes the problem of massless topological fermions interacting with massless vortex "Beryons," i.e., the quanta of the Berry gauge field a_μ, and is formally equivalent to the Euclidean quantum electrodynamics of massless Dirac fermions in 2+1 dimensions (QED_3). It, however, suffers from an intrinsic Dirac anisotropy by virtue of $v_F\neq v_\Delta$.

B. QED_3 Lagrangian for the pseudogap state

We are now in position to do some real calculations within our theory. Before we plunge into the algebra, however, we first apply some cosmetics: The Lagrangian (47) is not in the standard from as used in quantum electrodynamics where the matrices associated with the components of covariant derivatives form a Dirac algebra and mutually anticommute. In (47) the temporal derivative is associated with a unit matrix and it therefore commutes (rather than anticommutes) with σ_1 and σ_3 matrices associated with the spatial derivatives. These nonstandard commutation relations, however, lead to some rather unwieldy algebra. For this reason we manipulate the Lagrangian (47) into a slightly different form that is consistent with the usual field-theoretic notation. First, we combine each pair of antipodal (time-reversed) two-component spinors into one four-component spinor,

$$\Upsilon_1=\begin{pmatrix}\tilde{\Psi}_1\\ \tilde{\Psi}_{\bar{1}}\end{pmatrix},\quad \Upsilon_2=\begin{pmatrix}\tilde{\Psi}_2\\ \tilde{\Psi}_{\bar{2}}\end{pmatrix}. \tag{50}$$

Second, we define a new adjoint four-component spinor

$$\bar{\Upsilon}_l=-i\Upsilon_l^\dagger\gamma_0. \tag{51}$$

In terms of this new spinor the Lagrangian becomes

$$\mathcal{L}_D=\sum_{l=1,2}\bar{\Upsilon}_l\gamma_\mu D_\mu^{(l)}\Upsilon_l+\frac{1}{2}K_\mu(\partial\times a)_\mu^2, \tag{52}$$

with covariant derivatives

$$D_\mu^{(1)}=i[(\partial_\tau+ia_\tau),v_F(\partial_x+ia_x),v_\Delta(\partial_y+ia_y)],$$

$$D_\mu^{(2)}=i[(\partial_\tau+ia_\tau),v_F(\partial_y+ia_y),v_\Delta(\partial_x+ia_x)].$$

The 4×4 gamma matrices, defined as

$$\gamma_0=\begin{pmatrix}\sigma_2 & 0\\ 0 & -\sigma_2\end{pmatrix},\quad \gamma_1=\begin{pmatrix}\sigma_1 & 0\\ 0 & -\sigma_1\end{pmatrix},$$

$$\gamma_2=\begin{pmatrix}-\sigma_3 & 0\\ 0 & \sigma_3\end{pmatrix} \tag{53}$$

now form the usual Dirac algebra

$$\{\gamma_\mu,\gamma_\nu\}=2\delta_{\mu\nu} \tag{54}$$

and furthermore satisfy

$$\mathrm{Tr}(\gamma_\mu)=0,\quad \mathrm{Tr}(\gamma_\mu\gamma_\nu)=4\delta_{\mu\nu}. \tag{55}$$

The use of the adjoint spinor $\bar{\Upsilon}$ instead of the conventional $\Upsilon^\dagger$ is a purely formal device which will simplify calculations but does not alter the physical content of the theory. At the end of the calculation we have to remember to undo the transformation (51) by multiplying the $\langle\Upsilon(x)\bar{\Upsilon}(x')\rangle$ correlator by $i\gamma_0$ to obtain the physical correlator $\langle\Upsilon(x)\Upsilon^\dagger(x')\rangle$.

Next, to make the formalism simpler still we can eliminate the asymmetry between the two pairs of nodes by performing an internal SU(2) rotation at nodes 2, $\bar{2}$:

$$\Upsilon_2\to e^{-i(\pi/4)\gamma_0}\gamma_1\Upsilon_2, \tag{56}$$

leading to the anisotropic QED_3 Lagrangian

$$\mathcal{L}_D=\sum_{l=1,2}\bar{\Upsilon}_l v_\mu^{(l)}\gamma_\mu(i\partial_\mu-a_\mu)\Upsilon_l+\frac{1}{2}K_\mu(\partial\times a)_\mu^2, \tag{57}$$

with $v_\mu^{(1)}=(1,v_F,v_\Delta)$ and $v_\mu^{(2)}=(1,v_\Delta,v_F)$.

IV. SPECTRAL PROPERTIES OF TOPOLOGICAL FERMIONS AND PHYSICAL ELECTRONS IN QED_3

We shall start by considering the isotropic case $v_F=v_\Delta=1$, which although unphysical in the strictest sense is computationally much simpler and provides penetrating insights into the physics embodied by the QED_3 Lagrangian (57). After we have understood the isotropic case we will then be ready to tackle the calculation for the general case and will show that Dirac cone anisotropy does not modify the essential physics discussed here. To make contact with the standard literature on QED_3, we further consider a more general problem with *N pairs* of nodes described by the Lagrangian

$$\mathcal{L}_D=\sum_{l=1}^{N}\bar{\Upsilon}_l\gamma_\mu(i\partial_\mu-a_\mu)\Upsilon_l+\frac{1}{2}K_\mu(\partial\times a)_\mu^2. \tag{58}$$

For the basic problem of a single CuO_2 layer $N=2$. As we will show in the next section, *N* itself is variable and can be equal to 4 or 6 in bilayer and multilayer cuprates. Our analytic results can be viewed as arising from the formal $1/N$ expansion, although we expect them to be qualitatively (and even quantitatively) accurate even for $N=2$ as long as we are within the *symmetric* phase of QED_3 – the quantitative accuracy stems from a fortuitous conspiracy of small numerical prefactors.[38]

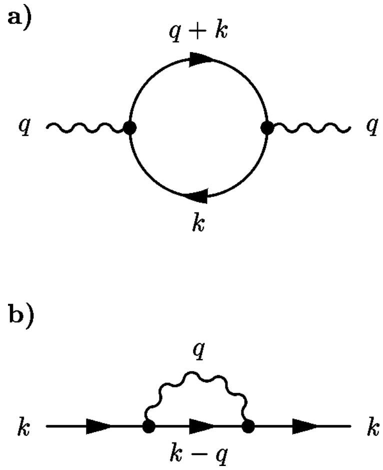

FIG. 3. One-loop Berryon polarization (a) and TF self-energy (b).

A. Berryon propagator

Ultimately, we are interested in the properties of physical electrons. To describe those we need to understand the properties of the electron-electron interaction mediated by the gauge field a_μ. To this end we proceed to calculate the Berry gauge field propagator by integrating out the fermion degrees of freedom from the Lagrangian (58). To one-loop order this corresponds to evaluating the vacuum polarization bubble, Fig. 3(a). Employing the standard rules for Feynman diagrams in the momentum space[39] the vacuum polarization reads

$$\Pi_{\mu\nu}(q)=N\int\frac{d^3k}{(2\pi)^3}\mathrm{Tr}[G_0(k)\gamma_\mu G_0(k+q)\gamma_\nu]. \quad (59)$$

Here $G_0(k)=k_\alpha\gamma_\alpha/k^2$ is the free Dirac Green function, $k=(k_0,\mathbf{k})$ denotes the Euclidean three-momentum, and the trace is performed over the γ matrices.

The integral in Eq. (59) is a standard one (see Appendix B for the details of computation) and the result is

$$\Pi_{\mu\nu}(q)=\frac{N}{8}|q|\left(\delta_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right), \quad (60)$$

where $|q|=\sqrt{q^2}$. The one-loop effective action for the Berry gauge field therefore becomes $(2\pi)^{-3}\int d^3q\mathcal{L}_B$ with

$$\mathcal{L}_B[a]=\left(\frac{N}{16}|q|+\frac{1}{2e^2}q^2\right)\left(\delta_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)a_\mu(q)a_\nu(-q). \quad (61)$$

At low energies and long wavelengths, $|q|/e^2\ll N/8$, the fermion polarization completely overwhelms the original Maxwell bare action term and the Berryon properties become universal. In particular, the coupling constant $1/e^2$ drops out at low energies and only reappears as the ultraviolet cutoff. Physically, the medium of massless Dirac fermions screens the long-range interactions mediated by a_μ. In QED_3 this screening is incomplete: the gauge field becomes stiffer by one power of q but still remains *massless*, in accordance with our general expectations. This anomalous stiffness of a_μ justifies the quadratic level expansion of $\mathcal{L}_0$, Eq. (45), and renders higher-order terms *irrelevant* in the RG sense. The theory therefore clears an important self-consistency check.

At low energies the fully dressed Berryon propagator is given as the inverse of the polarization,

$$D_{\mu\nu}(q)=\Pi^{-1}_{\mu\nu}(q). \quad (62)$$

In order to perform this inversion we have to fix the gauge. To this end we implement the usual gauge fixing procedure by replacing $q_\mu q_\nu/q^2\rightarrow(1-\xi^{-1})q_\mu q_\nu/q^2$ in Eq. (60). Here $\xi\geqslant 0$ parametrizes the orbit of all covariant gauges. For example, $\xi=0$ corresponds to the Lorentz gauge $k_\mu a_\mu(k)=0$ while $\xi=1$ corresponds to the Feynman gauge. Upon inversion we obtain the low-energy Berryon propagator

$$D_{\mu\nu}(q)=\frac{8}{|q|N}\left(\delta_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}(1-\xi)\right), \quad (63)$$

in agreement with previous authors.[41]

B. TF self-energy and propagator

The TF propagator is a gauge-dependent entity and one could therefore immediately object that as such it has no direct physical content and is of no interest. Such a viewpoint, while expressed frequently, is actually quite naive. The reality is that in gauge theories various gauge-variant objects can often be connected to physical gauge-invariant quantities when computed within a *particular* choice of gauge. In practice, a rather typical occurrence is that a gauge-invariant physical propagator is given by a hugely nonlocal form which is basically impossible to compute except in a judiciously chosen gauge where it is related to a much simpler, and therefore far easier to compute, gauge-variant propagator. Consequently, a gauge-variant propagator computed along a particular gauge orbit often contains relevant information about the true dynamics of a gauge theory—the trick is to know how to extract this information.

This general statement holds in the case of QED_3 as well. The TF propagator evaluated in an arbitrary covariant gauge parametrized by ξ contains useful information about the nature of the fermionic excitations of the system. We will show that its coupling to the massless gauge field destroys the usual Fermi liquid pole and results in the propagator displaying a Luttinger-like behavior, characterized by a small anomalous dimension. In the next subsection we argue that the physical electron propagator of our theory, which can be related to a particular gauge-invariant fermion propagator of QED_3, exhibits the same Luttinger-like behavior.

The lowest-order self-energy diagram is depicted in Fig. 3(b) and reads

$$\Sigma(k)=\int\frac{d^3q}{(2\pi)^3}D_{\mu\nu}(q)\gamma_\mu G_0(k+q)\gamma_\nu. \quad (64)$$

Again, the computation is rather straightforward (see Appendix B for details) and the most divergent part is

$$\Sigma(k)=\frac{4(2-3\xi)}{3\pi^2N}\not{k}\ln\left(\frac{\Lambda}{|k|}\right), \quad (65)$$

where we have introduced the Feynman "slash" notation $\not{k} = k_\mu \gamma_\mu$.

To the leading $1/N$ order, the inverse TF propagator is given by

$$G^{-1}(k) = \not{k}\left[1 + \eta \ln\left(\frac{\Lambda}{|k|}\right)\right], \tag{66}$$

with

$$\eta = -\frac{4(2-3\xi)}{3\pi^2 N}. \tag{67}$$

Higher-order contributions in $1/N$ will necessarily affect this result. Renormalization group arguments[42] and nonperturbative approaches[43] strongly suggest that Eq. (66) represents the start of a perturbative series that eventually resums into a power law:

$$G^{-1}(k) = \not{k}\left(\frac{\Lambda}{|k|}\right)^{\eta}. \tag{68}$$

This implies a real-space propagator of the form

$$G(r) = \Lambda^{-\eta}\frac{\not{r}}{r^{3+\eta}}. \tag{69}$$

Thus, the TF propagator exhibits a Luttinger-like algebraic singularity at small momenta, characterized by an anomalous exponent η. In the Lorentz gauge $(\xi=0)$ we find $\eta = -8/3\pi^2 N \simeq -0.13$, for $N=2$. This rather small numerical value for the anomalous dimension exponent (which is even considerably smaller for $N=4$ or $N=6$) indicates that the unraveling of the Fermi liquid pole in the original TF propagator brought about by its interaction with the massless Berry gauge field is in a certain sense "weak." Note also that η is *negative* in the Lorentz gauge while it becomes positive, $\eta = 4/3\pi^2 N \simeq 0.06$ for $N=2$, in the Feynman gauge $(\xi=1)$. The above results provide a strong indication that the physical, gauge-invariant fermion propagator also has a Luttinger-like form, characterized by a small and *positive* anomalous dimension.[15] We now show that this indeed is the case.

C. Physical electron propagator

Various spectroscopies on cuprates, such as angle-resolved photoemission spectroscopy (ARPES) and scanning tunneling microscopy (STM), as well as numerous optical and microwave techniques, all measure the spectral function of a real physical electron, not of TF's. Therefore, we are ultimately interested in computing the propagator of the physical electron in our theory,

$$G^{\text{elec}}(x-x') = \langle \Psi(x)\Psi^\dagger(x')\rangle, \tag{70}$$

where $\Psi(x)$ is the original electron field operator appearing in Eq. (1) (the reader should recall that this operator already contains high-energy renormalizations built in at the very beginning). If we define a matrix

$$\hat{\Omega}(x) = \begin{pmatrix} e^{-i\varphi_A(x)} & 0 \\ 0 & e^{i\varphi_B(x)} \end{pmatrix}, \tag{71}$$

we can write G^{elec} in terms of TF fields as

$$G^{\text{elec}}(x-x') = \langle \hat{\Omega}(x)\tilde{\Psi}(x)\tilde{\Psi}^\dagger(x')\hat{\Omega}^\dagger(x')\rangle. \tag{72}$$

Ordinary spectroscopies reflect the diagonal part of the electron propagator,

$$[G^{\text{elec}}(x-x')]_{11} = \langle e^{-i[\varphi_A(x)-\varphi_A(x')]}[\tilde{\Psi}(x)\tilde{\Psi}^\dagger(x')]_{11}\rangle, \tag{73}$$

and a similar expression for $[G^{\text{elec}}(x-x')]_{22}$. We may recast this in a more convenient form by writing the phase difference in the exponent as a line integral of a gradient along a *straight* line connecting x and x',

$$\varphi_A(x) - \varphi_A(x') \to -\int_x^{x'} \partial_\mu \varphi_A ds_\mu, \tag{74}$$

and then expressing the phase gradient in terms of the two gauge fields a_μ and v_μ:

$$[G^{\text{elec}}(x-x')]_{11} = \left\langle \exp\left(i\int_x^{x'}(v_\mu + a_\mu)ds_\mu\right) \times [\tilde{\Psi}(x)\tilde{\Psi}^\dagger(x')]_{11}\right\rangle. \tag{75}$$

This expression only involves the coarse-grained Doppler and Berry gauge fields and TF's, which are precisely the fields that enter our effective low-energy theory, and is therefore amenable to analysis. Note that Eq. (75) is only a long-distance approximation to the exact electron propagator (73) which is defined through discrete vortex variables entering via $\varphi_A(\varphi_B)$.

As discussed earlier, the Doppler gauge field v_μ is massive in both normal and superconducting phases and therefore its fluctuations will not affect the low-energy, long-wavelength properties of the electron propagator. We may thus remove it from the line integral, Eq. (75), and focus on the quantity

$$\mathcal{G}(x-x') = \left\langle \exp\left(i\int_x^{x'} a_\mu ds_\mu\right)\Upsilon(x)\bar{\Upsilon}(x')\right\rangle. \tag{76}$$

By considering the transformation properties of $\Upsilon(x)$ under the gauge transformations with respect to a_μ it is easy to verify that $\mathcal{G}(x-x')$ is *gauge-invariant*. This quantity therefore represents a gauge invariant propagator of QED$_3$ theory (52) and its knowledge allows us to reconstruct the diagonal components of the electron propagator by means of

$$[G^{\text{elec}}(x-x')]_{ii} = [i\gamma_0 \mathcal{G}(x-x')]_{ii}. \tag{77}$$

It seems natural to attempt to relate the components of the above gauge-invariant propagator to the physical electron propagator, since the latter by definition must be gauge invariant under the transformations of the internal gauge field a_μ.

The question arises as to how to evaluate the gauge-invariant propagator $\mathcal{G}(x-x')$. This turns out to be a nontrivial issue since, despite its pleasing manifest invariance under gauge and spacetime symmetries, Eq. (76) also exhibits a severe (linear) ultraviolet divergence arising from the straight line integral of the gauge field. This renders it ill defined in the absence of some proper regularization (see below for more details). Here we adopt the approach discussed by Brown[44,45] in the context of QED_4. Brown shows that the following relation exists between gauge invariant propagator $\tilde{\mathcal{G}}$ and gauge-dependent propagator G (Ref. 46) (see also Appendix C):

$$\tilde{\mathcal{G}}(r)=e^{F(r)}G(r), \tag{78}$$

where $r=x-x'$ and

$$F(r)=\frac{1}{2}\int d^3z\int d^3z' J_\mu(z)D_{\mu\nu}(z-z')J_\nu(z'), \tag{79}$$

with

$$J_\mu(z)=r_\mu\int_0^1 d\alpha\,\delta(z-x'-\alpha r) \tag{80}$$

representing the source term for the line integral in Eq. (76). In the above $G(r)$ and $D_{\mu\nu}(r)$ refer to the real-space gauge-dependent fermion and Berryon propagators, respectively, obtained by Fourier transforming the expressions (68) and (63). By definition, both are to be computed in *covariant* gauge. Brown's result, Eq. (78), is an explicit statement of the fact that one can construct two gauge-invariant propagators by using the line integral of the gauge fields $\mathcal{G}$ and $\tilde{\mathcal{G}}$. This is a rather general feature of Abelian gauge theories[45] and is easily generalized to QED_3. Here $\mathcal{G}$ and $\tilde{\mathcal{G}}$ can be formally related through a gauge transformation (see Appendix C) and one might think of $\tilde{\mathcal{G}}$ as representing suitably regularized $\mathcal{G}$.[47] Alternatively, we can simply think of $\tilde{\mathcal{G}}$ as being another QED_3 fermion propagator invariant under gauge and spacetime symmetries just like $\mathcal{G}$. We discuss the details pertaining to Eq. (78) in Appendix C.

To calculate $\tilde{\mathcal{G}}(r)$ from Eq. (78) we need to evaluate $F(r)$. We proceed by first performing the z, z' integrals to obtain

$$F(r)=\frac{1}{2}\int_0^1 d\alpha\int_0^1 d\beta\, r_\mu D_{\mu\nu}(r(\alpha-\beta))r_\nu. \tag{81}$$

By power counting the expression for $F(r)$ suffers from a linear UV divergence, reflecting the singular behavior of the gauge field line integral at short distances. This singularity is the main reason why direct computation of Eq. (76) is such a frustrating task, despite its deceivingly compact and elegant form. A typical scheme to regularize this linear UV divergence interferes with gauge invariance and corrupts the effort to extract the true physical part of Eq. (76) which we expect to be scale invariant. The advantage of Brown's approach (78) is twofold: it permits computation of a gauge-invariant propagator $\tilde{\mathcal{G}}$ in the *covariant* gauge where the UV divergence can be treated with the help of *dimensional regularization*, which respects the gauge invariance and preserves the long-wavelength, low-energy properties of the physical propagators, and the offending linear part of the UV divergence cancels out between numerator and denominator in a gauge-invariant manner (Appendix C). This allows us to extract a meaningful power-law behavior for $\tilde{\mathcal{G}}$. To take advantage of dimensional regularization we express $\mathcal{P}(r)\equiv r_\mu D_{\mu\nu}(r)r_\nu$ as a Fourier transform in d dimensions,

$$\mathcal{P}(r)=\frac{8}{N}\int\frac{d^dk}{(2\pi)^d}\frac{e^{ik\cdot r}}{k}\left[r^2-(1-\xi)\frac{(k\cdot r)^2}{k^2}\right], \tag{82}$$

and treat d as a continuous variable. This Fourier transform is evaluated in Appendix B, giving the result

$$\mathcal{P}(r)=\frac{4}{N}\frac{\Gamma\left(\frac{d-1}{2}\right)}{\pi^{(d+1)/2}}[1+(d-2)(1-\xi)]r^{3-d}. \tag{83}$$

Substituting $\mathcal{P}(r(\alpha-\beta))$ into Eq. (81) and performing the remaining integrals by means of

$$\int_0^1 d\alpha\int_0^1 d\beta|\alpha-\beta|^\zeta=\frac{2}{(\zeta+1)(\zeta+2)}, \tag{84}$$

we find, near $d=3$,

$$\begin{aligned}F(r)&=-\frac{4(2-\xi)}{N\pi^2}\lim_{d\to3}\left(\frac{r^{3-d}}{3-d}\right)\\&=-\frac{4(2-\xi)}{N\pi^2}\left[\ln(\Lambda r)+\frac{1}{3-d}\right]_{d\to3}.\end{aligned} \tag{85}$$

The UV divergence is now parametrized by the (r-independent) second term in the angular brackets. The leading long-distance behavior is contained in the logarithm, implying a power-law contribution $e^{F(r)}\propto r^{-4(2-\xi)/N\pi^2}$ to the gauge-invariant propagator $\tilde{\mathcal{G}}(r)$. Combining Eqs. (85) with (69) we obtain[48]

$$\tilde{\mathcal{G}}(r)=\Lambda^{-\eta'}\frac{\not{r}}{r^{3+\eta'}} \tag{86}$$

or, in momentum space,

$$\tilde{\mathcal{G}}(k)=\Lambda^{-\eta'}\frac{\not{k}}{k^{2-\eta'}}, \tag{87}$$

where the anomalous dimension exponent η' is given by

$$\eta'=\eta+\frac{4(2-\xi)}{N\pi^2}=\frac{16}{3\pi^2N}. \tag{88}$$

The last equation informs us that in the exponent η' the gauge fixing parameter ξ has canceled out and $\tilde{\mathcal{G}}(r)$ is indeed gauge invariant. We have thus verified, by explicit calculation to leading order in $1/N$, that Eq. (78) yields a gauge-invariant TF propagator which we can connect to the physical electron propagator by means of Eq. (77) with $\mathcal{G}\to\tilde{\mathcal{G}}$.

An interesting feature of the above result is that for $\xi=2$ we have $F(r)=0$. In the QED literature this is known as "Yennie's gauge" and its significance is that in this particular gauge the diagonal components of the TF propagator are directly equal to $\tilde{\mathcal{G}}(r)$. In Yennie's gauge one can therefore evaluate various electron observables in terms of the TF propagator without worrying about the exponential factors. This is just the situation we have anticipated in the previous subsection. One could also define an "anti-Yennie's gauge" (or a "nonlocal gauge" as it is known in the QED_3 literature) $\xi=\frac{2}{3}$, in which $\eta=0$ and the effect of the gauge field fluctuations on $\tilde{\mathcal{G}}(r)$ is contained entirely in $F(r)$. To leading order in $1/N$ this observation further solidifies the expectation that the leading logarithm in the self-energy indeed exponentiates and the low-energy, long-length-scale propagator behaves as a power law.

Another important feature to observe is that $\eta'>0$—the electron has acquired a *positive* anomalous dimension. The positivity of η' is mandatory from general considerations—once we perform the Euclidean rotation and obtain the real-time electron propagator the conditions of unitarity and causality of our original problem demand $\eta'>0$. This is also a physically sensible result implying that the interacting electron propagator (86) decays on long length scales *faster* than the free BdG electron propagator. Interaction mediated by the massless gauge field destabilizes the Fermi liquid pole in the original propagator and leads to a Luttinger liquid-like power-law fermionic correlator in the low-energy, long-wavelength limit—this is our algebraic Fermi liquid, a non-Fermi-liquid *symmetric* phase which governs the physics of the pseudogap state. The positivity of η' means that the interacting AFL propagator is *less coherent* than the free BdG electron propagator which is just what one expects on intuitive grounds. We therefore propose that $\tilde{\mathcal{G}}$ be identified as the true electron propagator in our theory.

The Luttinger-like electron propagator that follows from Eq. (87) leads to a characteristic asymmetry between the energy and momentum distribution curves (EDC and MDC) observed in ARPES experiments[49] (note that, within our theory, this behavior is limited to the pseudogap state)—the MDC is a very sharp Lorentzian close to the Fermi surface while the EDC is broad, reflecting the decoherence of physical electrons in the pseudogap state[15]). This decoherence is relatively weak with the explicit value of $\eta'\simeq 0.27$ for the case of an individual CuO_2 layer where $N=2$. In bilayer or multilayer systems N could be 4, 6, or even higher and η' is even smaller. The reason for this increase in N is the anisotropy of the tunneling matrix element between the constituent CuO_2 planes within a multilayer unit cell. This matrix element effectively vanishes near the nodes, along $(\pm\pi,\pm\pi)$ directions, but is appreciable elsewhere in the Brillouin zone.[50] The result is that low-energy BdG fermions on different constituent CuO_2 layers remain *decoupled* while the vortex excitations on these same layers are strongly *coupled* within the unit cell of a multilayer, since their coupling comes from an integral over the full Brillouin zone. This translates to a larger effective N in our QED_3, Eq. (58), and to a corresponding reduction in the anomalous dimension: $\eta'\simeq 0.13$ for $N=4$ (bilayer like YBCO) or $\eta'\simeq 0.09$ for $N=6$ (trilayer like $HgBa_2Ca_2Cu_3O_8$). The actual numerical value of the anomalous dimension exponent η' is the "fingerprint," a unique mathematical signature of the symmetric phase of QED_3 and therefore of the AFL state within the pseudogap regime of underdoped cuprates. Determining η' directly from experiments, either through various spectroscopies or transport measurements, would be a major step toward testing the theoretical ideas expressed in this paper.

The exponent η' of the physical electron propagator had come under much scrutiny as of late since several effective theories related to QED_3 emerged recently in condensed matter physics, in problems like Heisenberg antiferromagnets or spin liquids. While in each case the physical content of these multiple reincarnations of QED_3 differs completely from the one discussed in the present paper and from each other, the issue of the gauge-invariant QED_3 fermion propagator and the value of η' looms large in all these different contexts, for obvious reasons. In particular, η' has been calculated recently by Rantner and Wen[51,52] and also by Khveshchenko.[53] The former authors obtain $\eta'=-32/3\pi^2N$ (Ref. 54) by performing a calculation of $\mathcal{G}$ in the so-called axial gauge,[55] in which the line integral of the gauge field in Eq. (76) is taken to vanish for a particular direction in real space. A negative anomalous dimension $\eta'<0$ would imply that the interacting electron propagator is *more coherent* at long distances than the propagator of a free electron and this is prohibited on general grounds, as discussed above. For example, a negative η' produces a divergent electronic density of states and leads to unphysical singular behavior in various thermodynamic and transport quantities. Thus, the negative anomalous dimension for the physical electron should be, in our view, rejected out of hand. For the reader's benefit, we should stress that we believe that the calculations carried out in Refs. 51–53 are perfectly correct, in the sense that $\eta'=-32/3\pi^2N$ indeed follows from the algebra once we adopt the axial gauge regularization of Eq. (76) as implemented in Ref. 51 and 52 and perform the calculation with logarithmic accuracy; we have done such a calculation ourselves and have obtained the same result. Operationally, the problem with computation of $\mathcal{G}$, Eq. (76), is not in the algebra but resides in the physical interpretation of the obtained results. In the axial gauge regularization followed by the momentum space computation employed by the authors of Refs. 51 and 52, the negative value for η' arises from the treatment of spurious gauge singularities that are invariably introduced by writing the gauge boson propagator in the axial gauge.[55] These singularities then must be regularized in some way and this is done using an *ad hoc* principal value prescription for the momentum-space integrals. The problem with this prescription is that the momentum-space propagator in the axial gauge *does not exist*.[56] Axial gauge is an example of a singularly noncovariant gauge (like Coulomb, temporal, or similar gauges) and as such does not fully fix the gauge. The gauge transformations which are independent of the space-time variable x_1 singled out by the axial gauge but have arbitrary dependence on the remaining $d-1$ variables $(x_2,x_3,\ldots)$ are still allowed and cost no energy. Consequently, the gauge-variant two-point fermion propagator

$G(\mathbf{x},\mathbf{x}')$ must vanish whenever $x_2\neq x_2'$, $x_3\neq x_3'$, ...[56] The same problem resurfaces in a different form when one computes Eq. (76) directly in the covariant gauge:[57] the expectation value of the transverse part ($\xi=0$) of the line integral, $\langle \exp(i\int_x^{x'} a_\mu ds_\mu)\rangle$, decouples from the rest of the expression. In fact, the result $\eta'=-32/3\pi^2N$ is easily understood as arising from the TF propagator in the Lorentz gauge $G_{\xi=0}$ ($\eta_L=-8/3\pi^2N$) being made more coherent through simply being multiplied by the said expectation value of the transverse part of line integral ($\eta'=\eta_L+\eta_t$ where $\eta_t=-8/\pi^2N$ from the above computation of F, with $\xi=0$). Since $\langle \exp(i\int_x^{x'} a_\mu ds_\mu)\rangle_{\xi=0}$ is the expectation value of a phase factor, this is clearly a troubling result – such a multiplicative factor can make the full propagator only less coherent than $G_{\xi=0}$ as $|x-x'|\to\infty$. The problem is that the transverse part of the line integral is more divergent than just a simple $\ln r$ appearing in the exponent. When the dominant (linear) divergence is included, the full propagator $\mathcal{G}$ is exponentially suppressed ($\eta\to+\infty$) and cannot be computed without some physically motivated UV regularization scheme.[47,57]

We have also carried out our own calculation in the axial gauge using a different regularization scheme and obtained a different (negative) exponent $\eta'=-16/3\pi^2N$. We are thus forced to conclude that the axial gauge calculations yield values of η' that are regularization scheme dependent and are therefore inherently unreliable. By contrast no such ambiguities arise when employing Eq. (78) and the computation of $\tilde{\mathcal{G}}$ yields physically reasonable positive anomalous dimension given by Eq. (88).

We now discuss the off-diagonal (anomalous) components of the electron propagator. According to Eq. (71) we have

$$[G^{\rm elec}(x-x')]_{12}=\langle e^{-i[\varphi_A(x)+\varphi_B(x')]}[\tilde{\Psi}(x)\tilde{\Psi}^\dagger(x')]_{12}\rangle, \tag{89}$$

and a similar expression for $[G^{\rm elec}(x-x')]_{21}$. It is now less straightforward to interpret the phase factor in terms of our gauge fields a_μ and v_μ. One way to do this is to add and subtract $\varphi_A(x')$ and write

$$\begin{aligned}\varphi_A(x)+\varphi_B(x')&=[\varphi_A(x)-\varphi_A(x')]+[\varphi_A(x')+\varphi_B(x')]\\ &=-\int_x^{x'}(v_\mu+a_\mu)ds_\mu+2\int_0^{x'}v_\mu ds_\mu .\end{aligned} \tag{90}$$

The first term on the right-hand side is just like the one we encountered in our discussion of the diagonal propagator. The second term, however, involves a line integral of v_μ originating at an *arbitrary* fixed reference point in spacetime. In the nonsuperconducting phase fluctuations in v_μ will clearly drive any such term to zero, making the off-diagonal terms of the electron propagator vanish. This is consistent with our general expectation that the electron propagator does not exhibit anomalous off-diagonal correlations in the normal state.

Finally, pulling different strands together, we can write down the full electron Green function in the AFL phase of the pseudogap state. To make a connection with the notation prevalent in condensed matter physics we perform a Euclidean rotation $k_0\to i\omega$ in Eq. (87) and with help of Eq. (77) we obtain

$$G^{\rm elec}(\mathbf{k},\omega)=\Lambda^{-\eta'}\frac{\omega+\sigma_3\epsilon_{\mathbf{k}}}{[\epsilon_{\mathbf{k}}^2+\Delta_{\mathbf{k}}^2-\omega^2]^{1-\eta'/2}}, \tag{91}$$

where we have also restored the full electron dispersion $\epsilon_{\mathbf{k}}$ and the gap function $\Delta_{\mathbf{k}}$ with the understanding that the above form for the propagator is strictly valid only in the vicinity of the nodal point and close to the isotropic limit. We note that Eq. (91) implies an anomalous electron density of states

$$N(\omega)\sim\omega^{1+\eta'} \tag{92}$$

at low energies. It would be very interesting if such an anomalous electron density of states could be measured in tunneling experiments. Similarly, the Luttinger-like behavior of the propagator, Eq. (91), will be reflected in other physical observables.

V. EFFECTS OF DIRAC ANISOTROPY IN SYMMETRIC QED$_3$

It is natural to examine to what extent the theory is modified by the inclusion of the Dirac anisotropy, i.e., the finite difference in the Fermi velocity v_F and the gap velocity v_Δ. In actual materials the Dirac anisotropy $\alpha_D=v_F/v_\Delta$ decreases with decreasing doping from ~ 15 in the optimally doped to ~ 3 in the heavily underdoped samples.

There are two key issues: first, for a large enough number of Dirac fermion species N, how is the chirally symmetric infrared (IR) fixed point modified by the fact that $\alpha_D\neq 1$, and second, as we decrease N, does the chiral symmetry breaking occur at the same value of N as in the isotropic theory? In this section we address in detail the first issue and defer the discussion of the second one to part II.

We determine the effect of the Dirac anisotropy, marginal by power counting, by the perturbative renormalization group to first order in the large N expansion. To the leading order δ in the small anisotropy $\alpha_D=1+\delta$, we obtain the analytic value of the RG β_{α_D} function and find that it is proportional to δ; i.e., in the infrared the α_D decreases when $\delta>0$ and the anisotropic theory flows to the isotropic fixed point. On the other hand, when $\delta<0$, α_D increases in the IR and again the theory flows into the isotropic fixed point. These results hold even when anisotropy is not small as shown by numerical evaluation of the β function. Therefore, we conclude that the isotropic fixed point is stable against small anisotropy.

Furthermore, we show that in any covariant gauge renormalization of Σ due to the unphysical longitudinal degrees of freedom is *exactly* the same along any spacetime direction. Therefore the only contribution to the RG flow of anisotropy comes from the physical degrees of freedom and our results for β_{α_D} stated above are in fact gauge invariant.

A. Anisotropic QED$_3$

In the realm of condensed matter physics there is no Lorentz symmetry to safeguard the spacetime isotropy of the theory. Rather, the intrinsic Dirac anisotropy is always present since it ultimately arises from complicated microscopic interactions in the solid which eventually renormalize to band and pairing amplitude dispersion. Thus there is nothing to protect the difference in the Fermi velocity $v_F = \partial\epsilon_{\mathbf{k}}/\partial\mathbf{k}$ and the gap velocity $v_\Delta = \partial\Delta_{\mathbf{k}}/\partial\mathbf{k}$ from vanishing and, in fact, all HTS materials are anisotropic.

The value of α_D can be directly measured by ARPES, which is ultimately a ''high''-energy local probe of v_F and v_Δ. Since QED$_3$ is free on short distances, we can take the experimental values as the starting bare parameters of the field theory.

The pairing amplitude of the HTS cuprates has $d_{x^2-y^2}$ symmetry, and consequently there are four nodal points on the Fermi surface with Dirac dispersion around which we can linearize the theory. Note that the roles of x and y directions are interchanged between adjacent nodes. As before, we combine the four two-component Dirac spinors for the opposite (time-reversed) nodes into two four-components spinors and label them as $(1,\overline{1})$ and $(2,\overline{2})$ (see Fig. 2).

Thus, the two-point vertex function of the noninteracting theory for, say, $1\overline{1}$ fermions is

$$\Gamma^{(2)free}_{1\overline{1}} = \gamma_0 k_0 + v_F \gamma_1 k_1 + v_\Delta \gamma_2 k_2. \tag{93}$$

Therefore the corresponding noninteracting ''nodal'' Green functions are

$$G_0^n(k) = \frac{\sqrt{g^n}_{\mu\nu}\gamma_\mu k_\nu}{k_\mu g_{\mu\nu} k_\nu} \equiv \frac{\gamma_\mu^n k_\mu}{k_\mu g_{\mu\nu} k_\nu}. \tag{94}$$

Here we introduced the (diagonal) ''nodal'' metric $g_{\mu\nu}^{(n)}$: $g_{00}^{(1)} = g_{00}^{(2)} = 1$, $g_{11}^{(1)} = g_{22}^{(2)} = v_F^2$, $g_{22}^{(1)} = g_{11}^{(2)} = v_\Delta^2$, as well as the ''nodal'' γ matrices γ^n. In what follows we assume that both v_F and v_Δ are dimensionless and that eventually one of them can be chosen to be unity by an appropriate choice of the ''speed of light.''

Since $\alpha_D \neq 1$ breaks the Lorentz invariance of the theory and since it is the Lorentz invariance that protects the spacetime isotropy, we expect the β functions for α_D to acquire finite values. However, the theory still respects time reversal and parity and for N large enough the system is in the chirally symmetric phase. These symmetries force the fermion self-energy of the interacting theory to have the form

$$\Sigma_{1\overline{1}} = A(k_{1\overline{1}}, k_{2\overline{2}})(\gamma_0 k_0 + v_F \zeta_1 \gamma_1 k_1 + v_\Delta \zeta_2 \gamma_2 k_2), \tag{95}$$

where $k_{1\overline{1}} \equiv k_\mu g_{\mu\nu}^{(1)} k_\nu$ and $k_{2\overline{2}} \equiv k_\mu g_{\mu\nu}^{(2)} k_\nu$. The coefficients ζ_i are in general different from unity. Furthermore, there is a discrete symmetry which relates flavors $1,\overline{1}$ and $2,\overline{2}$ and the x and y directions in such a way that

$$\Sigma_{2\overline{2}} = A(k_{2\overline{2}}, k_{1\overline{1}})(\gamma_0 k_0 + v_\Delta \zeta_2 \gamma_1 k_1 + v_F \zeta_1 \gamma_2 k_2). \tag{96}$$

In the computation of the fermion self-energy, this discrete symmetry allows us to concentrate on a particular pair of nodes without any loss of generality.

B. Gauge field propagator

As discussed above in the isotropic case, the effect of vortex-antivortex fluctuations at $T=0$ on the fermions can, at large distances, be included by coupling the nodal fermions minimally to a fluctuating U(1) gauge field with a standard Maxwell action. Upon integrating out the fermions, the gauge field acquires a stiffness proportional to k, which is another way of saying that at the charged, chirally symmetric fixed point the gauge field has an anomalous dimension $\eta_A = 1$ (for a discussion of this point in bosonic QED, see Ref. 58).

We first proceed in the transverse gauge ($k_\mu a_\mu = 0$) which is in some sense the most physical one considering that the $\nabla \times \mathbf{a}$ is physically related to the vorticity, i.e., an intrinsically transverse quantity. We later extend our results to a general covariant gauge. To one-loop order the screening effects of the fermions on the gauge field are given by the polarization function

$$\Pi_{\mu\nu}(k) = \frac{N}{2} \sum_{n=1,2} \int \frac{d^3q}{(2\pi)^3} \mathrm{Tr}[G_0^n(q)\gamma_\mu^n G_0^n(q+k)\gamma_\nu^n], \tag{97}$$

where the index n denotes the fermion ''nodal'' flavor. The above expression can be evaluated straightforwardly by noting that it reduces to the isotropic $\Pi_{\mu\nu}(k)$ once the integrals are properly rescaled.[15] The result can be conveniently presented by taking advantage of the ''nodal'' metric $g_{\mu\nu}^n$ as

$$\Pi_{\mu\nu}(k) = \sum_n \frac{N}{16 v_F v_\Delta} \sqrt{k_\alpha g_{\alpha\beta}^n k_\beta} \left(g_{\mu\nu}^n - \frac{g_{\mu\rho}^n k_\rho g_{\nu\lambda}^n k_\lambda}{k_\alpha g_{\alpha\beta}^n k_\beta} \right). \tag{98}$$

Note that this expression is explicitly transverse, $k_\mu \Pi_{\mu\nu}(k) = \Pi_{\mu\nu}(k) k_\nu = 0$, and symmetric in its spacetime indices. It also properly reduces to the isotropic expression when $v_F = v_\Delta = 1$.

However, as opposed to the isotropic case, it is not quite as straightforward to determine the gauge field propagator $D_{\mu\nu}$. For example, as it stands the polarization matrix (98) is not invertible, which makes it necessary to introduce some gauge fixing conditions. In our case the direct inversion of the 3×3 matrix would obscure the analysis and, therefore, we choose to follow a more physical and notationally transparent line of reasoning which eventually leads to the correct expression for the gauge field propagator. Upon integrating out the fermions and expanding the effective action to one-loop order, we find that

$$\mathcal{L}_{eff}[a_\mu] = \tfrac{1}{2}(\Pi_{\mu\nu}^{(0)} + \Pi_{\mu\nu}) a_\mu a_\nu, \tag{99}$$

where the bare gauge field stiffness is

$$\Pi_{\mu\nu}^{(0)} = \frac{1}{e^2} k^2 \left(\delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} \right). \tag{100}$$

At this point we introduce the dual field b_μ which is related to a_μ as

$$b_\mu = \epsilon_{\mu\nu\lambda} q_\nu a_\lambda . \tag{101}$$

Physically, the b_μ field represents vorticity and we integrate over all possible vorticity configurations with the restriction that b_μ is transverse. We note that

$$\mathcal{L}[b_\mu] = \chi_0 b_0^2 + \chi_1 b_1^2 + \chi_2 b_2^2 , \tag{102}$$

where χ_μ's are functions of k_μ and upon a straightforward calculation they can be found to read

$$\chi_\mu = \frac{1}{2e^2} + \frac{N}{32 v_F v_\Delta} \sum_{n=1,2} \frac{g^n_{\nu\nu} g^n_{\lambda\lambda}}{\sqrt{k_\alpha g^n_{\alpha\beta} k_\beta}}, \tag{103}$$

where $\mu \neq \nu \neq \lambda \in \{0,1,2\}$. At low energies we can neglect the nondivergent bare stiffness and thus set $1/e^2 = 0$ in the above expression.

The expression (102) is manifestly gauge invariant and has the merit of not only being quadratic but also diagonal in the individual components of b_μ. Thus, integration over the vorticity (even with the restriction of transverse b_μ) is simple and we can easily determine the b_μ field correlation function

$$2\langle b_\mu b_\nu \rangle = \frac{\delta_{\mu\nu}}{\chi_\mu} - \frac{k_\mu k_\nu}{\chi_\mu \chi_\nu} \left(\sum_i \frac{k_i^2}{\chi_i} \right)^{-1} . \tag{104}$$

The repeated indices are not summed in the above expression. Note that, in addition to being transverse, $\langle b_\mu b_\nu \rangle$ is also symmetric in its spacetime indices.

It is now quite simple to determine the correlation function for the a_μ field and in the transverse gauge we obtain

$$D_{\mu\nu}(q) = \langle a_\mu a_\nu \rangle = \epsilon_{\mu ij} \epsilon_{\nu kl} \frac{q_i q_k}{q^4} \langle b_j b_l \rangle . \tag{105}$$

Using the transverse character of $\langle b_\mu b_\nu \rangle$ (which is independent of the gauge) the above expression can be further reduced to

$$D_{\mu\nu}(q) = \frac{1}{q^2} \left[\left(\delta_{\mu\nu} - \frac{q_\mu q_\nu}{q^2} \right) \langle b^2 \rangle - \langle b_\mu b_\nu \rangle \right] . \tag{106}$$

It can be easily checked that in the isotropic limit the expression (106) properly reduces to the results obtained in a different way.

We can further extend this result to include a general gauge by writing

$$D_{\mu\nu}(q) = \frac{1}{q^2} \left\{ \left[\delta_{\mu\nu} - \left(1 - \frac{\xi}{2} \right) \frac{q_\mu q_\nu}{q^2} \right] \langle b^2 \rangle - \langle b_\mu b_\nu \rangle \right\} , \tag{107}$$

where ξ is our continuous parametrization of the gauge fixing. This expression can be justified by the Faddeev-Popov type of procedure starting from the Lagrangian

$$\mathcal{L} = \frac{1}{2} \left(\Pi_{\mu\nu} + \frac{1}{\xi} \frac{2k^2}{\langle b^2 \rangle} \frac{k_\mu k_\nu}{k^2} \right) a_\mu a_\nu , \tag{108}$$

where the stiffness for the unphysical modes was judiciously chosen to scale as k in a particular combination of the physical scalars of the theory. Note that $\langle b^2 \rangle$ can be determined without ever considering gauge fixing terms. In this way, the extension of a transverse gauge $\xi = 0$ to a general covariant gauge is accomplished by a simple substitution $k_\mu k_\nu \to (1 - \xi/2) k_\mu k_\nu$. The expression (107) is our final result for the gauge field propagator in a covariant gauge.

C. TF self-energy

As discussed in Sec. IV, Σ is not gauge invariant in that it has an explicit dependence on the gauge fixing parameter. As we will show in this section (and more generally in Appendix D), the renormalization of Σ by unphysical longitudinal degrees of freedom does not depend on the spacetime direction: the term in Σ which is proportional to γ_0 is renormalized the same way by the gauge-dependent part of the action as the terms proportional to γ_1 and γ_2. Therefore, the only contribution to the RG flow of α_D comes from the physical degrees of freedom.

We denote the topological fermion self-energy at the node n by $\Sigma_n(q)$. Hence, to the leading order in a large-N expansion we have

$$\Sigma_n(q) = \int \frac{d^3k}{(2\pi)^3} \gamma^n_\mu G^n_0(q-k) \gamma^n_\nu D_{\mu\nu}(k) \tag{109}$$

or, explicitly,

$$\Sigma_n(q) = \int \frac{d^3k}{(2\pi)^3} \gamma^n_\mu \frac{(q-k)_\lambda \gamma^n_\lambda}{(q-k)_\mu g_{\mu\nu} (q-k)_\nu} \gamma^n_\nu D_{\mu\nu}(k), \tag{110}$$

where the gauge field propagator $D_{\mu\nu}$ is already screened by the nodal fermions (107). Using the fact that

$$\gamma_\mu \gamma_\lambda \gamma_\nu = i \epsilon_{\mu\lambda\nu} \gamma_5 \gamma_3 + \delta_{\mu\lambda} \gamma_\nu - \delta_{\mu\nu} \gamma_\lambda + \delta_{\lambda\nu} \gamma_\mu , \tag{111}$$

where $\mu,\nu,\lambda \in \{0,1,2\}$ and $\gamma_5 \equiv -i \gamma_0 \gamma_1 \gamma_2 \gamma_3$, we can easily see that

$$\gamma^n_\mu \gamma^n_\lambda \gamma^n_\nu D_{\mu\nu} = (2 g^n_{\lambda\mu} \gamma^n_\nu - \gamma^n_\lambda g^n_{\mu\nu}) D_{\mu\nu} , \tag{112}$$

where we used the symmetry of the gauge field propagator tensor $D_{\mu\nu}$. Thus,

$$\Sigma_n(q) = \int \frac{d^3k}{(2\pi)^3} \frac{(q-k)_\lambda (2 g^n_{\lambda\mu} \gamma^n_\nu - \gamma^n_\lambda g^n_{\mu\nu}) D_{\mu\nu}(k)}{(q-k)_\mu g^n_{\mu\nu} (q-k)_\nu}, \tag{113}$$

and as shown in the Appendix D at low energies this can be written as

$$\Sigma_n(q) = -\sum_\mu \eta^n_\mu (\gamma^n_\mu q_\mu) \ln \left(\frac{\Lambda}{\sqrt{q_\alpha g^n_{\alpha\beta} q_\beta}} \right) . \tag{114}$$

Here Λ is an upper cutoff and the coefficients η are functions of the bare anisotropy, which have been reduced to a quadrature (see Appendix D). It is straightforward, even if somewhat tedious, to show that in case of weak anisotropy ($v_F=1+\delta,\ v_\Delta=1$), to order δ^2,

$$\eta_0^{1\bar{1}}=-\frac{8}{3\pi^2N}\left(1-\frac{3}{2}\xi-\frac{1}{35}(40-7\xi)\delta^2\right), \quad (115)$$

$$\eta_1^{1\bar{1}}=-\frac{8}{3\pi^2N}\left(1-\frac{3}{2}\xi+\frac{6}{5}\delta-\frac{1}{35}(43-7\xi)\delta^2\right), \quad (116)$$

$$\eta_2^{1\bar{1}}=-\frac{8}{3\pi^2N}\left(1-\frac{3}{2}\xi-\frac{6}{5}\delta-\frac{1}{35}(1-7\xi)\delta^2\right). \quad (117)$$

In the isotropic limit ($v_F=v_\Delta=1$) we regain $\eta_\mu^n=-8(1-3/2\xi)/3\pi^2N$ as previously found by others.

D. Dirac anisotropy and its β function

Before plunging into any formal analysis, we wish to discuss some immediate observations regarding the RG flow of the anisotropy. Examining Eq. (114) it is clear that if $\eta_1^n=\eta_2^n$, then the anisotropy *does not* flow and remains equal to its bare value. That would mean that anisotropy is marginal and the theory flows into the anisotropic fixed point. In fact, such a theory would have a *critical line* of α_D. For this to happen, however, there would have to be a symmetry which would protect the equality $\eta_1^n=\eta_2^n$. For example, in isotropic QED$_3$ the symmetry which protects the equality of η's is the Lorentz invariance. In the case at hand, this symmetry is broken and therefore we expect that η_1^n will be different from η_2^n, suggesting that the anisotropy flows away from its bare value. If we start with $\alpha_D>1$ and find that $\eta_2^{1\bar{1}}>\eta_1^{1\bar{1}}$ at some scale $p<\Lambda$, we would conclude that the anisotropy is marginally irrelevant and decreases towards 1. On the other hand, if $\eta_2^{1\bar{1}}<\eta_1^{1\bar{1}}$, then anisotropy continues increasing beyond its bare value and the theory flows into a *critical point* with (in)finite anisotropy.

The issue is further complicated by the fact that η_μ^n is not a gauge-invariant quantity; i.e., it depends on the gauge fixing parameter ξ. The statement that, say, $\eta_1^n>\eta_2^n$ makes sense only if the ξ dependence of η_1^n and η_2^n is exactly the same; otherwise, we could choose a gauge in which the difference $\eta_2^n-\eta_1^n$ can have either sign. However, we see from Eqs. (115)–(117) that in fact the ξ dependence of all η's is indeed the same. Although it was explicitly demonstrated only to $\mathcal{O}(\delta^2)$, in Appendix D we show that it is in fact true to all orders of anisotropy for any choice of covariant gauge fixing. This fact provides the justification for our procedure. Now we supply the formal analysis reflecting the above discussion.

The renormalized two-point vertex function is related to the "bare" two-point vertex function via a fermion field rescaling Z_ψ as

$$\Gamma_R^{(2)}=Z_\psi\Gamma^{(2)}. \quad (118)$$

It is natural to demand that, for example, at nodes 1 and $\bar{1}$ at some renormalization scale p, $\Gamma_R^{(2)}(p)$ will have the form

$$\Gamma_R^{(2)}(p)=\gamma_0p_0+v_F^R\gamma_1p_1+v_\Delta^R\gamma_2p_2. \quad (119)$$

Thus, Eq. (119) corresponds to our renormalization condition through which we can eliminate the cutoff dependence and calculate the RG flows.

To the order of $1/N$ we can write

$$\Gamma_R^{(2)}(p)=Z_\psi\gamma_\mu^n p_\mu\left(1+\eta_\mu^n\ln\frac{\Lambda}{p}\right), \quad (120)$$

where we used the fermionic self-energy (114). Multiplying both sides by γ_0 and tracing the resulting expression determines the field strength renormalization

$$Z_\psi=\frac{1}{1+\eta_0^n\ln\frac{\Lambda}{p}}\approx 1-\eta_0^n\ln\frac{\Lambda}{p}. \quad (121)$$

We can now determine the renormalized Fermi and gap velocities

$$\frac{v_F^R}{v_F}\approx\left(1-\eta_0^{1\bar{1}}\ln\frac{\Lambda}{p}\right)\left(1+\eta_1^{1\bar{1}}\ln\frac{\Lambda}{p}\right)\approx 1-(\eta_0^{1\bar{1}}-\eta_1^{1\bar{1}})\ln\frac{\Lambda}{p} \quad (122)$$

and

$$\frac{v_\Delta^R}{v_\Delta}\approx\left(1-\eta_0^{1\bar{1}}\ln\frac{\Lambda}{p}\right)\left(1+\eta_2^{1\bar{1}}\ln\frac{\Lambda}{p}\right)\approx 1-(\eta_0^{1\bar{1}}-\eta_2^{1\bar{1}})\ln\frac{\Lambda}{p}. \quad (123)$$

The corresponding renormalized Dirac anisotropy is therefore

$$\alpha_D^R\equiv\frac{v_F^R}{v_\Delta^R}\approx\alpha_D\left(1-(\eta_2^{1\bar{1}}-\eta_1^{1\bar{1}})\ln\frac{\Lambda}{p}\right). \quad (124)$$

The RG beta function can now be determined:

$$\beta_{\alpha_D}=\frac{d\alpha_D^R}{d\ln p}=\alpha_D(\eta_2^{1\bar{1}}-\eta_1^{1\bar{1}}). \quad (125)$$

In the case of weak anisotropy ($v_F=1+\delta,\ v_\Delta=1$) the above expression can be determined analytically as an expansion in δ. Using Eqs. (116) and (117) we obtain

$$\beta_{\alpha_D}=\frac{8}{3\pi^2N}\left(\frac{6}{5}\delta(1+\delta)(2-\delta)+\mathcal{O}(\delta^3)\right). \quad (126)$$

Note that this expression is independent of the gauge fixing parameter ξ. For $0<\delta\ll 1$ the β function is positive which means that anisotropy decreases in the IR and thus anisotropic QED$_3$ scales to an isotropic QED$_3$. For $-1\ll\delta<0$ the β function is negative and in this case the anisotropy increases towards the fixed point $\alpha_D=1$, i.e., again towards the isotropic QED$_3$. Note that for $\delta>2$, $\beta<0$, which may na-

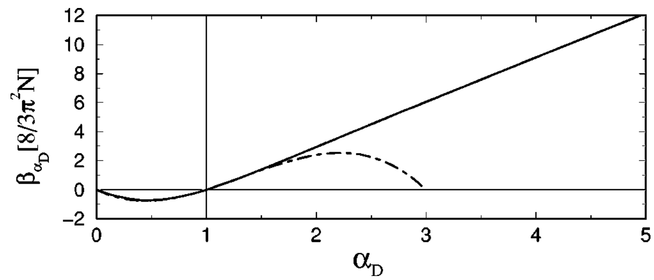

FIG. 4. The RG β function for the Dirac anisotropy in units of $8/3\pi^2N$. The solid line is the numerical integration of the quadrature in Eq. (D8) while the dash-dotted line is the analytical expansion around the small anisotropy [see Eq. (115)–(117)]. At $\alpha_D=1$, β_{α_D} crosses zero with positive slope, and therefore at large length scales anisotropic QED_3 scales to an isotropic theory.

ively indicate that there is a fixed point at $\delta=2$; this, however, cannot be trusted as it is outside the range of validity of the power expansion of η_μ. The numerical evaluation of the quadrature in Eq. (D8) shows that, apart from the isotropic fixed point and the unstable fixed point at $\alpha_D=0$, β_{α_D} does not vanish (see Fig. 4). This indicates that to leading order in the $1/N$ expansion, the theory flows into the isotropic fixed point.

VI. SUMMARY AND CONCLUSIONS

By appealing to the unique features of high-T_c cuprates—strong electron correlations, unconventional order parameter symmetry, and pronounced fluctuation effects—we argued in favor of the inverted approach to the problem of describing various thermodynamic phases appearing in the underdoped region. This inverted approach can be thought of as a ''Fermi liquid'' theory of the phase fluctuating *d*-wave superconductor where the role of the Fermi energy as the large energy scale of the problem is played by Δ, the amplitude of the pseudogap, which we assume to be predominantly of pairing nature. Under the umbrella of this pseudogap—inside the pairing protectorate—we identify the BdG quasiparticles as the relevant low-lying fermionic excitations of the theory and study their evolution under the effects of interactions mediated by vortex-antivortex fluctuations. By carefully treating these interactions we find that the low-energy effective theory for the quasiparticles inside the pairing protectorate is (2+1)-dimensional quantum electrodynamics (QED_3) with inherent spatial anisotropy, described by the Lagrangian $\mathcal{L}_D$ specified by Eqs. (47) and (57).

Within the superconducting state the gauge fields of the theory are massive by virtue of vortex defects being bound into finite loops or vortex-antivortex pairs. Such massive gauge fields produce only short-ranged interactions between our BdG quasiparticles and are therefore irrelevant: in the superconductor quasiparticles remain sharp in agreement with prevailing experimental data.[59] Loss of long-range superconducting order is brought about by unbinding the topological defects—vortex loops or vortex-antivortex pairs—via a Kosterlitz-Thouless-type transition and its quantum cousin. Remarkably, this is accompanied by the Berry gauge field becoming massless. Such a massless gauge field mediates long-range interactions between the fermions and becomes a relevant perturbation. Exactly what the consequence of this relevant interaction is depends on the number of fermion species N in the problem. For cuprates we argued that $N=2n_{\mathrm{CuO}}$ where n_{CuO} is the number of CuO_2 layers per unit cell. If $N<N_c\simeq 3$,[20] the interactions cause a spontaneous opening of a gap for the fermionic excitations at $T=0$, via the mechanism of chiral symmetry breaking in QED_3.[19] The formation of the gap corresponds to the onset of AF SDW instability[21,22] which must be considered as a progenitor of the Mott-Hubbard-Néel antiferromagnet at half-filling. If, on the other hand, $N>N_c$ as will be the case in bilayer or trilayer materials, the theory remains in its chirally symmetric nonsuperconducting phase even as $T\to 0$ and AF order arises from within such a state only upon further underdoping (Fig. 1). We call this symmetric state of QED_3 an algebraic Fermi liquid. In both cases the AFL controls the low-temperature, low-energy behavior of the pseudogap state and in this sense assumes the role played by Fermi liquid theory in conventional metals and superconductors. In the AFL the quasiparticle pole is replaced by a branch cut—the quasiparticle is no longer sharp—and the gauge-invariant electron propagator acquires a Luttinger-like form, Eq. (87), with positive anomalous dimension $\eta'=16/3\pi^2N$. To our knowledge this is one of the very few cases where a non-Fermi-liquid nature of the excitations has been demonstrated in dimension greater than 1 in the absence of disorder or a magnetic field. This Luttinger-like behavior of the AFL will manifest itself in an anomalous power-law functional form of many physical properties of the system.

Dirac anisotropy, i.e., the fact that $v_F\neq v_\Delta$, plays an important role in the cuprates where the ratio $\alpha_D=v_F/v_\Delta$ in most materials ranges between 3 and 15–20. Such anisotropy is nontrivial as it cannot be rescaled and it significantly complicates any calculation within the theory. Using the perturbative renormalization group theory we have shown that anisotropic QED_3 flows back into an isotropic stable fixed point. This means that for weak anisotropy at long length scales the universal properties of the theory are identical to those of the simple isotropic case. It remains to be seen what the properties of the theory are at intermediate length scales when the anisotropy is strong.

Our theory of the pseudogap state gets its inspiration and builds on the ideas originally articulated by Emery and Kivelson in Ref. 8 and by Randeria and collaborators.[60] These ideas were later explored and extended in various directions by others.[14,17,18,61–63] These approaches share a common philosophical platform of assuming that the pseudogap is primarily due to pairing in the p-p channel and in the underdoped regime superconducting long-range coherence is destroyed by the phase fluctuations. The essential differences in physics lie in the implementation of these ideas. Balents *et al.*,[14] for instance, argue for a true separation of spin and charge[32] in their ''nodal liquid'' phase which then gives way to an unconventional antiferromagnet ''AF*'' with deconfined spin-1/2 excitations. Our theory also shares considerable formal similarities with the SU(2) gauge theories of Kim, Lee, and Wen[41] and works by Aitchinson and

Mavromatos,[19] in that their low-energy effective theory is also related to QED$_3$. In those cases, however, the physical contents of the theory—the identity of its excitations—are totally different.

Finally there is a class of theories where the pseudogap is assumed to be due to some competing order, usually in the p-h channel. This includes SO(5),[64] d-density-wave,[65] and various other competing orders.[66] At the present time we believe that experimental evidence favors the pairing origin of the pseudogap; however, the evidence is far from conclusive and experiments can be found to support virtually any aspect of the above-mentioned theories. We believe, therefore, that the road ahead necessitates making specific predictions based on controlled and well-defined approaches. The QED$_3$ theory of the pairing pseudogap, as presented in this paper, starts from a remarkably simple set of assumptions, and via manipulations that are controlled in the sense of $1/N$ expansion, arrives at nontrivial consequences, including the algebraic Fermi liquid and the insulating antiferromagnet. Another notable feature of this theory is that it is fully calculable: having the explicit form of the low-energy effective action, free of constraints and uncertainties, physical observables can be computed through a systematic procedure.[67]

ACKNOWLEDGMENTS

The authors are grateful to P. W. Anderson, A. Auerbach, T. Appelquist, G. Baskaran, A. Cohen, Ø. Fischer, M. P. A. Fisher, B. Giovannini, S. M. Girvin, L. P. Gor'kov, V. P. Gusynin, I. F. Herbut, C. Kallin, D. V. Khveshchenko, S. A. Kivelson, R. B. Laughlin, A. J. Millis, H. Monien, N. P. Ong, W. Rantner, M. Reenders, S. Sachdev, T. Senthil, D. E. Sheehy, A. Sudbø, A.-M. Tremblay, P. B. Wiegmann, A. Zee, and S. C. Zhang for helpful discussions and correspondence. This research was supported in part by NSERC (M.F.) and the NSF Grant No. PHY99-07949 (ITP) and No. DMR00-94981 (Z.T., O.V.).

APPENDIX A: JACOBIAN $\mathcal{L}_0$

Here we derive the explicit form of the "Jacobian" $\mathcal{L}_0$ for two cases of interest: (i) the thermal vortex-antivortex fluctuations in 2D layers and (ii) the spacetime vortex loop excitations relevant for low temperatures ($T \ll T^*$) deep in the underdoped regime.

1. Two-dimensional thermal vortex-antivortex fluctuations

In order to perform specific computations we have to adopt a model for vortex-antivortex excitations. We will use a 2D Coulomb gas picture of vortex-antivortex plasma. In this model (anti)vortices are either pointlike objects or are assumed to have a small hard-disk radius of size of the coherence length ξ_0 which emulates the core region. As long as $\xi_0 \ll n^{-(1/2)}$, where $n = n_v + n_a$ is the average density of vortex defects, the two models lead to very similar results and both undergo a vortex-antivortex pair unbinding transition of the Kosterlitz-Thouless variety.

Above the transition we have

$$\exp\left[-\beta\int d^2r\mathcal{L}_0\right] = 2^{-N_l}\sum_{A,B}\int \mathcal{D}\varphi(\mathbf{r})\,\delta[\nabla\times\mathbf{v} - \tfrac{1}{2}\nabla\times(\nabla\varphi_A + \nabla\varphi_B)]\,\delta[\nabla\times\mathbf{a} - \tfrac{1}{2}\nabla\times(\nabla\varphi_A - \nabla\varphi_B)]. \tag{A1}$$

The phase $\varphi(\mathbf{r})$ is due solely to vortices and we can rewrite Eq. (A1) as

$$\sum_{N_v, N_a}\frac{2^{-N_l}}{N_v!N_a!}\sum_{A,B}\prod_i^{N_v}\int d^2r_i\prod_j^{N_a}\int d^2r_j e^{-\beta E_c(N_v+N_a)}\delta\left(\rho_v(\mathbf{r}) - \sum_i^{N_v}\delta(\mathbf{r}-\mathbf{r}_i)\right)\delta\left(\rho_a(\mathbf{r}) - \sum_j^{N_a}\delta(\mathbf{r}-\mathbf{r}_j)\right)$$
$$\times\delta\left(b(\mathbf{r}) - \pi\sum_i^{N_v^A}\delta(\mathbf{r}-\mathbf{r}_i^A) + \pi\sum_i^{N_v^B}\delta(\mathbf{r}-\mathbf{r}_i^B) + \pi\sum_j^{N_a^A}\delta(\mathbf{r}-\mathbf{r}_j^A) - \pi\sum_i^{N_a^B}\delta(\mathbf{r}-\mathbf{r}_j^B)\right). \tag{A2}$$

Here $N_v(N_a)$ is the number of free vortices [antivortices], $N_l = N_v + N_a$, $\mathbf{r}_i$ [$\mathbf{r}_j$] are vortex (antivortex) coordinates, and $\rho_v(\mathbf{r})$ [$\rho_a(\mathbf{r})$] are the corresponding densities. $b(\mathbf{r}) = (\nabla\times a(\mathbf{r}))_z = \pi(\rho_v^A - \rho_v^B - \rho_a^A + \rho_a^B)$ and E_c is the core energy which we have absorbed into $\mathcal{L}_0$ for convenience. We now express the above δ functions as functional integrals over three new fields $d_v(\mathbf{r})$, $d_a(\mathbf{r})$, and $\kappa(\mathbf{r})$:

$$\sum_{N_v, N_a}\frac{2^{-N_l}}{N_v!N_a!}\sum_{A,B}\prod_i^{N_v}\int d^2r_i\prod_j^{N_a}\int d^2r_j e^{-\beta E_c(N_v+N_a)}\int\mathcal{D}d_v\mathcal{D}d_a\mathcal{D}\kappa\exp\left\{i\int d^2r d_v\left(\rho_v(\mathbf{r}) - \sum_i^{N_v}\delta(\mathbf{r}-\mathbf{r}_i)\right)\right.$$
$$+ i\int d^2r d_a\left(\rho_a(\mathbf{r}) - \sum_j^{N_a}\delta(\mathbf{r}-\mathbf{r}_j)\right) + i\int d^2r\,\kappa\left[b(\mathbf{r}) - \pi\sum_i^{N_v^A}\delta(\mathbf{r}-\mathbf{r}_i^A)\right.$$
$$\left.\left. + \pi\sum_i^{N_v^B}\delta(\mathbf{r}-\mathbf{r}_i^B) + \pi\sum_j^{N_a^A}\delta(\mathbf{r}-\mathbf{r}_j^A) - \pi\sum_i^{N_a^B}\delta(\mathbf{r}-\mathbf{r}_j^B)\right]\right\}. \tag{A3}$$

The integration over δ functions in the exponential is easily performed and the summation over $A(B)$ labels can be carried out explicitly to obtain

$$\int \mathcal{D}d_v\mathcal{D}d_a\mathcal{D}\kappa \exp\left[i\int d^2r(d_v\rho_v+d_a\rho_a+\kappa b)\right]\sum_{N_v,N_a}\frac{e^{-\beta E_c N_v}}{N_v!}\prod_i^{N_v}\int d^2r_i \exp[-id_v(\mathbf{r}_i)]$$
$$\times\cos[\pi\kappa(\mathbf{r}_i)]\frac{e^{-\beta E_c N_a}}{N_a!}\prod_j^{N_a}\int d^2r_j\exp[-id_a(\mathbf{r}_j)]\cos[\pi\kappa(\mathbf{r}_j)]. \quad \text{(A4)}$$

In the thermodynamic limit the sum (A4) is dominated by $N_{v(a)}=\langle N_{v(a)}\rangle$, where $\langle N_{v(a)}\rangle$ is the average number of free (anti)vortices determined by solving the full problem. Furthermore, as $\langle N_{v(a)}\rangle\to\infty$ in the thermodynamic limit the integration over $d_v(\mathbf{r})$, $d_a(\mathbf{r})$, and $\kappa(\mathbf{r})$ can be performed in the saddle-point approximation, leading to the following saddle-point equations:

$$-\rho_v(\mathbf{r})+\langle N_v\rangle\Omega_v(\mathbf{r})=0, \quad \text{(A5)}$$

$$-\rho_a(\mathbf{r})+\langle N_a\rangle\Omega_a(\mathbf{r})=0, \quad \text{(A6)}$$

$$-b(\mathbf{r})+[\langle N_v\rangle\Omega_v(\mathbf{r})+\langle N_a\rangle\Omega_a(\mathbf{r})]\pi\tanh[\pi\kappa(\mathbf{r})]=0, \quad \text{(A7)}$$

with

$$\Omega_m(\mathbf{r})=\frac{e^{d_m(\mathbf{r})}\cosh[\pi\kappa(\mathbf{r})]}{\int d^2r' e^{d_m(\mathbf{r}')}\cosh[\pi\kappa(\mathbf{r}')]},\quad m=a,v.$$

Equations (A5)–(A7) follow from functional derivatives of Eq. (A4) with respect to $d_v(\mathbf{r})$, $d_a(\mathbf{r})$, and $\kappa(\mathbf{r})$, respectively. We have also built in the fact that the saddle-point solutions occur at $d_v\to id_v$, $d_a\to id_a$, and $\kappa\to i\kappa$.

The saddle-point equations (A5)–(A7) can be solved exactly, leading to

$$d_v(\mathbf{r})=\ln\rho_v(\mathbf{r})-\ln\cosh[\pi\kappa(\mathbf{r})], \quad \text{(A8)}$$

$$d_a(\mathbf{r})=\ln\rho_a(\mathbf{r})-\ln\cosh[\pi\kappa(\mathbf{r})], \quad \text{(A9)}$$

where

$$\kappa(\mathbf{r})=\frac{1}{\pi}\tanh^{-1}\left[\frac{b(\mathbf{r})}{\pi[\rho_v(\mathbf{r})+\rho_a(\mathbf{r})]}\right]. \quad \text{(A10)}$$

Inserting Eqs. (A8)–(A10) back into Eq. (A4) finally gives the entropic part of $\mathcal{L}_0/T$:

$$\rho_v\ln\rho_v+\rho_a\ln\rho_a-\frac{1}{\pi}(\nabla\times\mathbf{a})_z\tanh^{-1}\left[\frac{(\nabla\times\mathbf{a})_z}{\pi(\rho_v+\rho_a)}\right]$$
$$+(\rho_v+\rho_a)\ln\cosh\tanh^{-1}\left[\frac{(\nabla\times\mathbf{a})_z}{\pi(\rho_v+\rho_a)}\right], \quad \text{(A11)}$$

where $\rho_{v(a)}(\mathbf{r})$ are densities of *free* (anti)vortices. We display $\mathcal{L}_0$ in this form to make contact with familiar physics: the first two terms in Eq. (A11) are the entropic contribution of free (anti)vortices and the Doppler gauge field $\nabla\times\mathbf{v}\to\pi(\rho_v-\rho_a)$ ($\langle\nabla\times\mathbf{v}\rangle=0$). The last two terms encode the "Berry phase" physics of topological frustration. Note that the "Berry" magnetic field $b=(\nabla\times\mathbf{a})_z$ couples directly only to the *total density* of vortex defects $\rho_v+\rho_a$ and is insensitive to the vortex charge. This is a reflection of the Z_2 symmetry of the original problem defined on discrete (i.e., not coarse-grained) vortices. We write $\rho_{v(a)}(\mathbf{r})=\langle\rho_{v(a)}\rangle+\delta\rho_{v(a)}(\mathbf{r})$ and expand Eq. (A11) to leading order in $\delta\rho_{v(a)}$ and $\nabla\times\mathbf{a}$:

$$\mathcal{L}_0/T\to(\nabla\times\mathbf{v})^2/(2\pi^2n_l)+(\nabla\times\mathbf{a})^2/(2\pi^2n_l), \quad \text{(A12)}$$

where $n_l=\langle\rho_v\rangle+\langle\rho_a\rangle$ is the average density of free vortex defects. Both $\mathbf{v}$ and $\mathbf{a}$ have a Maxwellian *bare* stiffness and are *massless* in the normal state. As one approaches T_c, $n_l\sim\xi_{sc}^{-2}\to0$, where $\xi_{sc}(x,T)$ is the superconducting correlation length, and $\mathbf{v}$ and $\mathbf{a}$ become *massive* (see the main text).

2. Quantum fluctuations of (2+1)D vortex loops

The expression for $\mathcal{L}_0[j_\mu]$ given by Eq. (39) follows directly once the system contains unbound vortex loops in its ground state and thus can respond to the external perturbation A_μ^{ext} over arbitrary large distances in (2+1)-dimensional spacetime. This is already clear at intuitive level if we just think of the geometry of infinite versus finite loops and the fact that only unbound loops allow vorticity fluctuations, described by $\langle j_\mu(q)j_\nu(-q)\rangle$, to proceed unhindered. Still, it is useful to derive Eq. (39) and its consequences belabored in Sec. II within an explicit model for vortex loop fluctuations.

Here we consider fluctuating vortex loops in continuous (2+1)D spacetime and compute $\mathcal{L}_0[j_\mu]$ using a duality map to the relativistic Bose superfluid.[68] We again start by using our model of a large-gap *s*-wave superconductor which enables us to neglect a_μ in the fermion action and integrate over it in the expression for $\mathcal{L}_0[v_\mu,a_\mu]$, Eq. (11). This gives

$$e^{-\int d^3x\mathcal{L}_0}=\sum_{N=0}^{\infty}\frac{1}{N!}\prod_{l=1}^{N}\oint\mathcal{D}x_l[s_l]\delta(j_\mu(x)-n_\mu(x))e^{-\tilde{S}}, \quad \text{(A13)}$$

where

$$\tilde{S}=\sum_{l=1}^{N}S_0\oint ds_l+\frac{1}{2}\sum_{l,l'=1}^{N}\oint ds_l\oint ds_{l'}g(|x_l[s_l]-x_{l'}[s_{l'}]|) \quad \text{(A14)}$$

and

$$(\partial\times\partial\varphi(x))_\mu=2\pi n_\mu(x)$$
$$=2\pi\sum_l^N \oint_L dx_{l\mu}\delta(x-x_l[s_l]). \quad \text{(A15)}$$

In the above equations N is the number of loops, s_l is the Schwinger proper time (or "proper length") of loop l, S_0 is the action per unit length associated with motion of vortex cores in (2+1)-dimensional spacetime (in an analogous 3D model this would be ε_c/T, where ε_c is the core line energy), $g(|x_l[s_l]-x_{l'}[s_{l'}]|)$ is the short-range penalty for core overlap, and L denotes a line integral. We kept our practice of including core terms independent of vorticity in $\mathcal{L}_0$. Note that vortex loops must be periodic along τ, reflecting the original periodicity of $\varphi(\mathbf{r},\tau)$.

We can think of vortex loops as world line trajectories of some relativistic charged (complex) bosons (charged since the loops have two orientations) in two spatial dimensions. $\mathcal{L}_0$ without the δ function describes the vacuum Lagrangian of such a theory, with vortex loops representing particle-antiparticle virtual creation and annihilation processes. The duality map is based on the following relation between the Green function of free charged bosons and a gas of free oriented loops:

$$G(x)=\langle\phi(0)\phi^*(x)\rangle$$
$$=\int\frac{d^3k}{(2\pi)^3}\frac{e^{ik\cdot x}}{k^2+m_d^2}=\int_0^\infty ds e^{-sm_d^2}\int\frac{d^3k}{(2\pi)^3}e^{ik\cdot x-sk^2}$$
$$=\int_0^\infty ds e^{-sm_d^2}\left(\frac{1}{4\pi s}\right)^{3/2}e^{-x^2/4s}, \quad \text{(A16)}$$

where m_d is the mass of the complex dual field $\phi(x)$. Within the Feynman path integral representation we can write

$$\left(\frac{1}{4\pi s}\right)^{3/2}e^{(-1/4)x^2/s}$$
$$=\int_{x(0)=0}^{x(s)=x}\mathcal{D}x(s')\exp\left[-\frac{1}{4}\int_0^s ds'\dot{x}^2(s')\right], \quad \text{(A17)}$$

where $\dot{x}\equiv dx/ds$. Furthermore, by simple integration Eq. (A16) can be manipulated into

$$\mathrm{Tr}[\ln(-\partial^2+m_d^2)]=-\int_0^\infty\frac{ds}{s}e^{-sm_d^2}\int\frac{d^3k}{(2\pi)^3}e^{-sk^2}$$
$$=-\int_0^\infty\frac{ds}{s}e^{-sm_d^2}\oint\mathcal{D}x(s')$$
$$\times\exp\left[-\frac{1}{4}\int_0^s ds'\dot{x}^2(s')\right], \quad \text{(A18)}$$

where the path integral now runs over closed loops $[x(s)=x(0)]$. In the dual theory this can be reexpressed as

$$\mathrm{Tr}[\ln(-\partial^2+m_d^2)]=\int\frac{d^3k}{(2\pi)^3}\ln(k^2+m_d^2). \quad \text{(A19)}$$

Combining Eqs. (A18) and (A19) and using

$$\mathrm{Tr}[\ln(-\partial^2+m_d^2)]=\ln\mathrm{Det}(-\partial^2+m_d^2)\equiv\mathcal{W} \quad \text{(A20)}$$

finally leads to

$$e^{-\mathcal{W}}=\sum_{N=0}^\infty\frac{1}{N!}\prod_{l=1}^N\left[\int_0^\infty\frac{ds_l}{s_l}e^{-s_lm_d^2}\oint\mathcal{D}x(s_l')\right]$$
$$\times\exp\left[-\frac{1}{4}\sum_{l=1}^N\int_0^{s_l}ds_l'\dot{x}^2(s_l')\right]. \quad \text{(A21)}$$

This is nothing else but the partition function of the free loop gas. The size of loops is regulated by m_d. As $m_d\to 0$ the average loop size diverges. On the other hand, through $\mathcal{W}$, Eq. (A20), we can also think of Eq. (A21) as the partition function of the free bosonic theory.

To exploit this equivalence further we write

$$Z_d=\int\mathcal{D}\phi^*\mathcal{D}\phi\exp\left(-\int d^3x\mathcal{L}_d\right), \quad \text{(A22)}$$

where

$$\mathcal{L}_d=|\partial\phi|^2+m_d^2|\phi|^2+\frac{g}{2}|\phi|^4, \quad \text{(A23)}$$

and argue that Z_d describes vortex loops with short range interactions in Eq. (A13). This can be easily demonstrated by decoupling $|\phi|^4$ through a Hubbard-Stratonovich transformation and retracing the above steps. The reader is referred to the book by Kleinert for further details of the above duality mapping.[68]

We can now rewrite the δ function in Eq. (A13) as

$$\delta(j_\mu(x)-n_\mu(x))\to\int\mathcal{D}\kappa_\mu\exp\left(i\int d^3x\kappa_\mu(j_\mu-n_\mu)\right) \quad \text{(A24)}$$

and observe that in the above language of Feynman path integrals in proper time $i\int d^3x\kappa_\mu n_\mu$, Eq. (A15), assumes the meaning of a particle current three-vector n_μ coupled to a three-vector potential κ_μ. Employing the same arguments that led to Eq. (A23) we now have

$$\mathcal{L}_d[\kappa_\mu]=|(\partial-i\kappa)\phi|^2+m_d^2|\phi|^2+\frac{g}{2}|\phi|^4, \quad \text{(A25)}$$

which leads to

$$\exp\left(-\int d^3x\mathcal{L}_0[j_\mu]\right)\to\int\mathcal{D}\phi^*\mathcal{D}\phi\mathcal{D}\kappa_\mu$$
$$\times\exp\left(-\int d^3x(-i\kappa_\mu j_\mu+\mathcal{L}_d[\kappa_\mu])\right). \quad \text{(A26)}$$

In the pseudogap state vortex loop unbinding causes loss of superconducting phase coherence. In the dual language, the appearance of such infinite loops as $m_d \to 0$ implies superfluidity in the system of charged bosons described by $\mathcal{L}_d$, Eq. (A23). The dual off-diagonal long range order (ODLRO) in $\langle \phi(x)\phi^*(x')\rangle$ means that there are world line trajectories that connect points x and x' over infinite spacetime distances—these infinite world lines are nothing else but unbound vortex paths in this "vortex loop condensate." Thus, the dual and the true superconducting ODLRO's are opposite sides of the same coin.

$\mathcal{L}_d[\kappa_\mu]$ is the Lagrangian of this dual superfluid in the presence of the external vector potential κ_μ. In the ordered phase, the response of the system is just the dual version of the Meissner effect. Consequently, upon functional integration over ϕ we are allowed to write

$$\exp\left(-\int d^3x \mathcal{L}_0[j_\mu]\right) = \int \mathcal{D}\kappa_\mu \exp\left(-\int d^3x(-i\kappa_\mu j_\mu + M_d^2 \kappa_\mu \kappa_\mu)\right), \tag{A27}$$

where $M_d^2 = |\langle \phi \rangle|^2$, with $\langle \phi \rangle$ being the dual order parameter. The remaining functional integration over κ_μ finally results in

$$\mathcal{L}_0[j_\mu] \to \frac{j_\mu j_\mu}{4|\langle \phi \rangle|^2}. \tag{A28}$$

We have tacitly assumed that the system of loops is isotropic. The intrinsic anisotropy of the (2+1)D theory is easily reinstated and Eq. (A28) becomes Eq. (39) of the main text.

It is now time to recall that we are interested in a *d*-wave superconductor. This means we must restore a_μ to the problem. To accomplish this we engage in a bit of thievery: imagine now that it was v_μ whose coupling to fermions was negligible and we could integrate over it in Eq. (11). We would then be left with only the δ function containing a_μ. Actually, we can compute such $\mathcal{L}_0[b_\mu/\pi]$, where $\partial \times \partial a = b$, without any additional work. Note that $\mathcal{L}_0$ contains only vorticity-independent terms. We can equally well proclaim that it is the $A(B)$ labels that determine the true orientation of our loops while the actual vorticity is simply a gauge label—in essence, v_μ and a_μ trade places. After the same algebra as before we obtain

$$\mathcal{L}_0[b_\mu/\pi] \to \frac{b_\mu b_\mu}{4\pi^2 |\langle \phi \rangle|^2}, \tag{A29}$$

which is just the Maxwell action for a_μ.

Of course, this simple argument that led to Eq. (A29) is illegal. We cannot just forget v_μ. If we did, we would have no right to coarse-grain a_μ to begin with and would have to face up to its purely Z_2 character (see Sec. II). Still, the above reasoning does illustrate that the Maxwellian stiffness of a_μ follows the same pattern as that of v_μ: both are determined by the order parameter $\langle \phi \rangle$ of condensed *dual* superfluid. Thus, we can write the correct form of $\mathcal{L}_0$, with both v_μ and a_μ fully included into our accounting, as

$$\mathcal{L}_0[v_\mu, a_\mu] \to \frac{(\partial\times v)_\mu(\partial\times v)_\mu}{4\pi^2|\langle \phi \rangle|^2} + \frac{(\partial\times a)_\mu(\partial\times a)_\mu}{4\pi^2|\langle \phi \rangle|^2}, \tag{A30}$$

where our ignorance is now stored in computing the actual value of $\langle \phi \rangle$ from the original parameters of the d-wave superconductor problem. With anisotropy restored this is precisely Eq. (45) of Sec. II.

APPENDIX B: FEYNMAN INTEGRALS IN QED$_3$

Many of the integrals encountered here are considered standard in particle physics. Since the techniques involved are not as common in condensed matter physics, we provide some of the technical details in this appendix. A more in-depth discussion can be found in many field theory textbooks.[39]

1. Vacuum polarization bubble

The vacuum polarization, Eq. (59), can be written more explicitly as

$$\Pi_{\mu\nu}(q) = 2N \,\mathrm{Tr}[\gamma_\alpha \gamma_\mu \gamma_\beta \gamma_\nu] I_{\alpha\beta}(q), \tag{B1}$$

with

$$I_{\alpha\beta}(q) = \int \frac{d^3k}{(2\pi)^3} \frac{k_\alpha(k_\beta + q_\beta)}{k^2(k+q)^2}. \tag{B2}$$

The integrals of this type are most easily evaluated by employing the Feynman parametrization.[39] This consists in combining denominators using the formula

$$\frac{1}{A^a B^b} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 dx \frac{x^{a-1}(1-x)^{b-1}}{[xA + (1-x)B]^{a+b}}, \tag{B3}$$

valid for any positive real numbers a, b, A, B. Setting $A = k^2$ and $B = (k+q)^2$ allows us to rewrite Eq. (B2) as

$$I_{\alpha\beta}(q) = \int_0^1 dx \int \frac{d^3k}{(2\pi)^3} \frac{k_\alpha(k_\beta + q_\beta)}{[(k+(1-x)q)^2 + x(1-x)q^2]^2}. \tag{B4}$$

We now shift the integration variable $k \to k - (1-x)q$ to obtain

$$I_{\alpha\beta}(q) = \int_0^1 dx \int \frac{d^3k}{(2\pi)^3} \frac{k_\alpha k_\beta - x(1-x)q_\alpha q_\beta}{[k^2 + x(1-x)q^2]^2}, \tag{B5}$$

where we have dropped terms odd in k which vanish by symmetry upon integration. We now notice that the $k_\alpha k_\beta$ term will only contribute if $\alpha = \beta$ and we can therefore replace it in Eq. (B5) by $\frac{1}{3}\delta_{\alpha\beta}k^2$. With this replacement the angular integrals are trivial and we have

$$I_{\alpha\beta}(q)=\frac{1}{2\pi^2}\int_0^1 dx\int_0^\infty dk k^2\frac{\frac{1}{3}\delta_{\alpha\beta}k^2-x(1-x)q_\alpha q_\beta}{[k^2+x(1-x)q^2]^2}. \tag{B6}$$

The only remaining difficulty stems from the fact that the integral formally diverges at the upper bound. This divergence is an artifact of our linearization of the fermionic spectrum which breaks down for energies approaching the superconducting gap value. It is therefore completely legitimate to introduce an ultraviolet cutoff. Such UV cutoffs, however, tend to interfere with gauge invariance, the preservation of which is crucial in this computation. A more physical way of regularizing the integral Eq. (B6) is to recall that the gauge field must remain massless, i.e., $\Pi_{\mu\nu}(q\to 0)=0$. To enforce this property we write Eq. (B1) as

$$\Pi_{\mu\nu}(q)=2N\,\mathrm{Tr}[\gamma_\alpha\gamma_\mu\gamma_\beta\gamma_\nu][I_{\alpha\beta}(q)-I_{\alpha\beta}(0)], \tag{B7}$$

and we see that proper regularization of Eq. (B6) involves subtracting the value of the integral at $q=0$. The remaining integral is convergent and elementary; explicit evaluation gives

$$I_{\alpha\beta}(q)-I_{\alpha\beta}(0)=-\frac{|q|}{64}\left(\delta_{\alpha\beta}+\frac{q_\alpha q_\beta}{q^2}\right). \tag{B8}$$

Inserting this in Eq. (B7) and working out the trace using Eqs. (54) and (55) we find the result (60). Identical result can be obtained using dimensional regularization.

2. TF self-energy: Lorentz gauge

For simplicity we evaluate the self-energy (64) in the Lorentz gauge ($\xi=0$). Extension to arbitrary covariant gauge is trivial. Equation (64) can be written as

$$\Sigma_L(k)=-\frac{2}{\pi^3 N}\gamma_\mu I_\mu(k), \tag{B9}$$

with

$$I_\mu(k)=\int d^3q\, q_\mu\frac{q\cdot(k+q)}{|q|^3(k+q)^2}, \tag{B10}$$

where we used an identity

$$\gamma_\mu\gamma_\alpha\gamma_\nu\left(\delta_{\mu\nu}-\frac{q_\mu q_\nu}{q^2}\right)=-2\not{q}\frac{q_\alpha}{q^2}.$$

Since the only three-vector available is k, clearly the vector integral $I_\mu(k)$ can only have components in the k_μ direction, $I_\mu(k)=\mathcal{C}(k)k_\mu/k^2$. By forming a scalar product $k_\mu I_\mu(k)$ we obtain

$$\mathcal{C}(k)=k_\mu I_\mu(k)=\int d^3q\frac{(q\cdot k)(q\cdot k+q^2)}{|q|^3(k+q)^2}. \tag{B11}$$

Combining the denominators using Eq. (B3) and following the same steps as in the computation of polarization bubble above we obtain

$$\mathcal{C}(k)=\frac{3}{2}\int_0^1 dx\sqrt{x}\int d^3q \times\frac{(2x-1)(k\cdot q)^2-(1-x)k^2q^2+x(1-x)^2k^4}{[q^2+x(1-x)k^2]^{5/2}}. \tag{B12}$$

The angular integrals are trivial and after introducing an ultraviolet cutoff Λ and rescaling the integration variable by $|k|$ the integral becomes

$$\mathcal{C}(k)=2\pi k^2\int_0^1 dx\sqrt{x}\int_0^{\Lambda/|k|}dq \times\frac{(5x-4)q^4+3x(1-x)^2q^2}{[q^2+x(1-x)]^{5/2}}. \tag{B13}$$

The remaining integrals are elementary. Isolating the leading infrared divergent term we obtain

$$\mathcal{C}(k)\to-\frac{4\pi}{3}k^2\ln\frac{\Lambda}{|k|}. \tag{B14}$$

Substituting this into Eq. (B9) we get the self-energy (65).

3. Dimensionally regularized gauge field line integral

To evaluate $\mathcal{P}(r)$ specified by Eq. (82) we write it as a sum of two contributions,

$$\mathcal{P}(r)=\frac{8r^2}{(2\pi)^dN}[I_1(r)-(1-\xi)I_2(r)], \tag{B15}$$

where

$$I_1(r)=\int d^dk\frac{e^{ik\cdot r}}{k}, \tag{B16}$$

$$I_2(r)=\int d^dk\frac{e^{ik\cdot r}}{k}\frac{(k\cdot r)^2}{k^2r^2}. \tag{B17}$$

We first consider $I_1(r)$. It is convenient to exponentiate the denominator by use of the formula

$$\frac{1}{A^a}=\frac{1}{\Gamma(a)}\int_0^\infty ds s^{a-1}e^{-sA}. \tag{B18}$$

Taking $A=k^2$, $a=\frac{1}{2}$ we have

$$I_1(r)=\frac{1}{\sqrt{\pi}}\int_0^\infty\frac{ds}{\sqrt{s}}\int d^dk e^{-sk^2+ik\cdot r}. \tag{B19}$$

The k integral is now easy to evaluate by completing the square, shifting the integration variable, and making use of $\int d^dk\exp(-sk^2)=(\pi/s)^{d/2}$. We thus obtain

$$I_1(r)=\pi^{(d-1)/2}\int_0^\infty ds s^{-(d+1)/2}e^{-r^2/4s}. \tag{B20}$$

The substitution $t=r^2/4s$ transforms Eq. (B20) to an integral of the type shown in Eq. (B18) and the result reads

$$I_1(r)=(4\pi)^{(d-1)/2}\Gamma\left(\frac{d-1}{2}\right)r^{1-d}. \tag{B21}$$

Using the same procedure we obtain

$$I_2(r)=-(4\pi)^{(d-1)/2}(d-2)\Gamma\left(\frac{d-1}{2}\right)r^{1-d}. \tag{B22}$$

Substituting Eqs. (B21) and (B22) back into Eq. (B15) we obtain the result quoted in the text, Eq. (83).

APPENDIX C: BROWN'S RELATION FOR A GAUGE-INVARIANT PROPAGATOR

To derive Brown's relation, Eq. (78), it is useful to introduce a fermion propagator for a *fixed* configuration of gauge field a_μ,

$$G[x,x';a]=\frac{1}{z[a]}\int \mathcal{D}[\overline{\Upsilon},\Upsilon]\Upsilon(x)\overline{\Upsilon}(x')e^{-S[\overline{\Upsilon},\Upsilon,a]}, \tag{C1}$$

where $S[\overline{\Upsilon},\Upsilon,a]=\int d^3r\mathcal{L}_D$ and $\mathcal{L}_D$ is the QED_3 Lagrangian given by Eq. (58), $z[a]=\int\mathcal{D}[\overline{\Upsilon},\Upsilon]e^{-S[\overline{\Upsilon},\Upsilon,a]}$. It is easy to see that $G[x,x';a]$ solves the Dirac equation

$$\gamma_\mu(i\partial_\mu-a_\mu)G[x,x';a]=\delta(x-x'). \tag{C2}$$

Furthermore, in terms of $G[x,x';a]$ the TF propagator can be written as

$$G(x-x')=\frac{1}{Z}\int \mathcal{D}a\, G[x,x';a]e^{-S_B[a]}, \tag{C3}$$

where $S_B=\int d^3x\mathcal{L}_B$ is the effective action [for our case, to one loop $\mathcal{L}_B$ is given by Eq. (61)] and $Z=\int\mathcal{D}a\, e^{-S_B[a]}$. We now introduce a *gauge-invariant* analog of $G[x,x';a]$,

$$\mathcal{G}[x,x';a]=\exp\left(-i\int_x^{x'}a\cdot dr\right)G[x,x';a], \tag{C4}$$

where the integral in the exponent is taken along the straight line connecting spacetime points x' and x. The gauge-invariant TF propagator defined by Eq. (76) is then given by an expression analogous to Eq. (C3),

$$\mathcal{G}(x-x')=\frac{1}{Z}\int \mathcal{D}a\, \mathcal{G}[x,x';a]e^{-S_B[a]}. \tag{C5}$$

Our task is to relate Brown's propagator $\tilde{\mathcal{G}}(x-x')$ to $G(x-x')$, Eq. (78). To this end we rewrite Eq. (C3) as follows:

$$\begin{aligned}G(x-x')&=\frac{1}{Z}\int \mathcal{D}a\, \mathcal{G}[x,x';a]e^{iJ\cdot a-S_B[a]}\\&=\left(\frac{\tilde{Z}}{Z}\right)\left[\frac{1}{\tilde{Z}}\int \mathcal{D}a\, \mathcal{G}[x,x';a]e^{-\tilde{S}_B[a]}\right]\\&=\left(\frac{\tilde{Z}}{Z}\right)\tilde{\mathcal{G}}(x-x'),\end{aligned} \tag{C6}$$

where $\tilde{S}_B[a]=S_B[a]-iJ\cdot a$, $\tilde{Z}=\int\mathcal{D}a\ \ e^{-\tilde{S}_B[a]}$, and the last equality in Eq. (C6) should be taken as a definition of $\tilde{\mathcal{G}}(x-x')$. The source term J is defined by Eq. (80) and $J\cdot a$ is shorthand for $\int d^3rJ(r)\cdot a(r)\equiv\int_x^{x'}a\cdot dr$. Note that the linear UV divergence of phase factor in $\tilde{Z}$ cancels out between the numerator and denominator, Eq. (C6).

We now observe that

$$\frac{\tilde{Z}}{Z}=\frac{1}{Z}\int \mathcal{D}a\, e^{iJ\cdot a}e^{-S_B[a]}=\langle e^{iJ\cdot a}\rangle\equiv e^{-F(x-x')}, \tag{C7}$$

with F defined by Eq. (79). Substituting this into Eq. (C6) we have, Eq. (78),

$$G(x-x')=e^{-F(x-x')}\tilde{\mathcal{G}}(x-x'). \tag{C8}$$

To complete this part it remains to address the relation between $\tilde{\mathcal{G}}(x-x')$ and $\mathcal{G}(x-x')$. To this end we notice that both $\mathcal{G}$ and $\tilde{\mathcal{G}}$ are *gauge invariant*: the former by construction and the latter by the following simple observation. Since both $\mathcal{G}[x,x';a]$ and $S_B[a]$ depend on the transverse part of a, any dependence on the longitudinal part of a comes through the $iJ\cdot a$ term and therefore identically cancels between the numerator and $\tilde{Z}$ in the denominator. Therefore, once defined by Eq. (C6) as a ratio of two gauge-variant objects computed in the same covariant gauge, $\tilde{\mathcal{G}}$ takes on life of its own and is equal to a fully gauge-invariant quantity. We may further rewrite $\tilde{\mathcal{G}}$ as such a ratio in arbitrary noncovariant linear gauge; specifically, we may chose the axial gauge in which the $iJ\cdot a$ vanishes. In this gauge $\tilde{\mathcal{G}}$ is exactly equal to $\mathcal{G}$. Since they are both gauge invariant, they must be equal in arbitrary gauge. The reader should be cautioned, however, that this last equality is a formal one since it involves manipulations of gauge-dependent quantities which are ill defined in the absence of some specific regularization.

APPENDIX D: FEYNMAN INTEGRALS FOR THE ANISOTROPIC CASE: SELF-ENERGY

As shown in Sec. V, to first order in the $1/N$ expansion, the topological fermion self-energy can be reduced to

$$\Sigma_n(q)=\int\frac{d^3k}{(2\pi)^3}\frac{(q-k)_\lambda(2g^n_{\lambda\mu}\gamma^n_\nu-\gamma^n_\lambda g^n_{\mu\nu})Dx_{\mu\nu}(k)}{(q-k)_\mu g^n_{\mu\nu}(q-k)_\nu}, \tag{D1}$$

where $D_{\mu\nu}(q)$ is the screened gauge field propagator evaluated in Sec. III. In order to perform the radial integral, we rescale the momenta as

$$K_\mu = \sqrt{g^n}_{\mu\nu} k_\nu, \quad Q_\mu = \sqrt{g^n}_{\mu\nu} q_\nu \tag{D2}$$

and obtain

$$\Sigma_n(q) = \int \frac{d^3K}{(2\pi)^3} \times \frac{(Q-K)_\lambda (2\sqrt{g^n}_{\lambda\mu}\gamma_\nu^n - \gamma_\lambda g^n_{\mu\nu}) D_{\mu\nu}\left(\frac{K_\nu}{\sqrt{g^n}_{\mu\nu}}\right)}{v_F v_\Delta (Q-K)^2}. \tag{D3}$$

At low energies we can neglect the contribution from the bare field stiffness ρ in the gauge propagator (see Sec. III) and the resulting $D_{\mu\nu}(q)$ exhibits $1/q$ scaling

$$D_{\mu\nu}\left(\frac{K_\nu}{\sqrt{g^n}_{\mu\nu}}\right) = \frac{F_{\mu\nu}(\theta,\phi)}{|K|}. \tag{D4}$$

Now we can explicitly integrate over the magnitude of the rescaled momentum K by introducing an upper cutoff Λ and in leading order we find that

$$\int_0^\Lambda dK \frac{K^2 (Q-K)_\lambda}{K(Q-K)^2} = -\Lambda \hat{K}_\lambda + \ln\left(\frac{\Lambda}{Q}\right)(Q_\lambda - 2\hat{K}_\lambda \hat{K}\cdot Q), \tag{D5}$$

where $\hat{K} = (\cos\theta,\ \sin\theta\cos\phi,\ \sin\theta\sin\phi)$. Since $\hat{K}_\mu$ is odd under inversion while $F_{\mu\nu}$ is even, it is not difficult to see that the term proportional to Λ vanishes upon angular integration. Thus

$$\Sigma_n(q) = \int \frac{d\Omega}{(2\pi)^3 v_F v_\Delta}(Q_\lambda - 2\hat{K}_\lambda \hat{K}\cdot Q) \times (2\sqrt{g^n}_{\lambda\mu}\gamma_\nu^n - \gamma_\lambda g^n_{\mu\nu}) F_{\mu\nu}(\theta,\phi) \ln\left(\frac{\Lambda}{Q}\right). \tag{D6}$$

Using the fact that the diagonal elements of $F_{\mu\nu}(\theta,\phi)$ are even under parity, while the off-diagonal elements are odd under parity, the above expression can be further simplified to

$$\Sigma_n(q) = -\sum_\mu (\gamma_\mu^n q_\mu \eta_\mu^n) \ln\left(\frac{\Lambda}{\sqrt{q_\alpha g^n_{\alpha\beta} q_\beta}}\right), \tag{D7}$$

where the coefficients η_μ^n are pure numbers depending on the bare anisotropy and are thereby reduced to a quadrature:

$$\eta_\mu^n = \int \frac{d\Omega}{v_F v_\Delta (2\pi)^3}\left[(2\hat{K}_\mu \hat{K}_\mu - 1)\left(2 g^n_{\mu\mu} F_{\mu\mu} - \sum_\nu g^n_{\nu\nu} F_{\nu\nu}\right) + \sum_{\nu\neq\mu} 4\hat{K}_\mu \hat{K}_\nu \sqrt{g^n}_{\mu\mu}\sqrt{g^n}_{\nu\nu} F_{\mu\nu}\right]. \tag{D8}$$

Repeated indices are not summed in the above expression, unless explicitly indicated. In the case of weak anisotropy ($v_F = 1+\epsilon, v_\Delta = 1$) we can show that Eq. (D8) reduces to Eqs. (115)–(117).

Consider now the effect of the (covariant) gauge fixing term on η_μ. Let us define the part of $F_{\mu\nu}$ which depends on the gauge fixing parameter ξ as $F^{(\xi)}_{\mu\nu}$. The general form of this term is $F^{(\xi)}_{\mu\nu} = \xi\ k_\mu k_\nu f(k)$ where $f(k)$ is a scalar function of all three components of k_μ; $f(k)$ does in general depend on the anisotropy. Upon rescaling with the nodal metric [see Eq. (D2)], we have

$$F^{(\xi)}_{\mu\nu} = \xi \frac{1}{\sqrt{g_{\mu\mu}}} K_\mu \frac{1}{\sqrt{g_{\nu\nu}}} K_\nu \tilde{f}(K),$$

where $\tilde{f}(K)$ is the corresponding scalar function of K_μ. Substituting $F^{(\xi)}_{\mu\nu}$ into the Eq. (D8) we find

$$\eta_\mu^\xi = \xi \int \frac{d\Omega\, \tilde{f}(K)}{v_F v_\Delta (2\pi)^3}\left[(2\hat{K}_\mu \hat{K}_\mu - 1)\left(2\hat{K}_\mu \hat{K}_\mu - \sum_\nu \hat{K}_\nu \hat{K}_\nu\right) + \sum_{\nu\neq\mu} 4\hat{K}_\mu \hat{K}_\nu \hat{K}_\mu \hat{K}_\nu\right], \tag{D9}$$

where η_μ^ξ is the part of η_μ which comes entirely from the gauge fixing term. Using the fact that $\Sigma_\nu \hat{K}_\nu \hat{K}_\nu = 1$, it is a matter of simple algebra to show that

$$\eta_\mu^\xi = \xi \int \frac{d\Omega\, \tilde{f}(K)}{v_F v_\Delta (2\pi)^3}; \tag{D10}$$

i.e., the dependence on the index μ drops out. That means that the renormalization of η_μ due to the unphysical longitudinal modes is exactly the same for all of its components. Therefore, the difference in η_1 and η_2, which is related to the RG flow of the Dirac anisotropy, comes entirely from the physical modes and is a gauge-independent quantity. Note that this statement does not depend on the choice of covariant gauge, i.e., on the exact form of the function f, only on the fact that the gauge is covariant.

[1] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957).
[2] G. M. Eliashberg, Sov. Phys. JETP **11**, 696 (1960); Y. Nambu, Phys. Rev. **117**, 648 (1960).
[3] W. Kohn and J. M. Luttinger, Phys. Rev. Lett. **15**, 524 (1965).
[4] P. W. Anderson, *The Theory of Superconductivity in the High-T_c Cuprates* (Princeton University Press, Princeton, 1997).
[5] P. Nozieres and D. Pines, *Theory of Quantum Liquids*, (Perseus

Books, Cambridge, MA, 1999), Vol. I.

[6] We use the term "BCS-Eliashberg-FL orthodoxy" in a narrow sense here to denote the theory of phonon-mediated pairing in conventional *s*-wave superconductors. A broader definition of the term could include various approaches based on spin fluctuations in nearly antiferromagnetic FL [P. Monthoux *et al.*, Phys. Rev. Lett. **67**, 3448 (1991)] which can lead to *d*-wave superconductivity.

[7] C. C. Tsuei and J. R. Kirtley, Rev. Mod. Phys. **72**, 969 (2000).

[8] V. J. Emery and S. A. Kivelson, Nature (London) **374**, 434 (1995).

[9] M. Tinkham, *Introduction to Superconductivity* (Krieger, 1975).

[10] Z. A. Xu *et al.*, Nature (London) **406**, 486 (2000); Y. Wang *et al.*, cond-mat/0108242 (unpublished).

[11] J. Corson *et al.*, Nature (London) **398**, 221 (1999).

[12] B. W. Hoogenboom *et al.*, Phys. Rev. B **62**, 9179 (2000).

[13] T. Nakano *et al.*, J. Phys. Soc. Jpn. **67**, 2622 (1998).

[14] L. Balents, M. P. A. Fisher, and C. Nayak, Phys. Rev. B **60**, 1654 (1999).

[15] M. Franz and Z. Tešanović, Phys. Rev. Lett. **87**, 257003 (2001).

[16] G. E. Volovik, JETP Lett. **58**, 469 (1993).

[17] M. Franz and A. J. Millis, Phys. Rev. B **58**, 14 572 (1998).

[18] H. J. Kwon and A. T. Dorsey, Phys. Rev. B **59**, 6438 (1999).

[19] R. D. Pisarski, Phys. Rev. D **29**, 2423 (1984); T. W. Appelquist *et al.*, *ibid.* **33**, 3704 (1986); M. R. Pennington and D. Walsh, Phys. Lett. B **253**, 246 (1991); I. J. R. Aitchison and N. E. Mavromatos, Phys. Rev. B **53**, 9321 (1996); P. Maris, Phys. Rev. D **54**, 4049 (1996).

[20] T. Appelquist, A. G. Cohen, and M. Schmaltz, Phys. Rev. D **60**, 045003 (1999). These authors argue that on general grounds $N_c<\frac{3}{2}$.

[21] I. F. Herbut, Phys. Rev. Lett. **88**, 047006 (2002).

[22] Z. Tešanović, O. Vafek, and M. Franz, Phys. Rev. B **65**, 180511 (2002).

[23] M. Franz and Z. Tešanović, Phys. Rev. Lett. **84**, 554 (2000).

[24] O. Vafek, A. Melikyan, M. Franz, and Z. Tešanović, Phys. Rev. B **63**, 134509 (2001); O. Vafek, A. Melikyan, and Z. Tešanović, Phys. Rev. B **64**, 224508 (2001).

[25] X.-G. Wen and P. A. Lee, Phys. Rev. Lett. **76**, 503 (1996); P. A. Lee, N. Nagaosa, T. K. Ng, and X.-G. Wen, Phys. Rev. B **57**, 6003 (1998).

[26] L. Marinelli, B. I. Halperin, and S. H. Simon, Phys. Rev. B **62**, 3488 (2000).

[27] The labels "compact" and "noncompact" gauge fields pasted on a_μ and v_μ have the following precise meaning: let us for a moment forget about definitions (7) and (8) and simply treat a_μ and v_μ appearing in Eq. (6) as two completely *independent* gauge fields. We further introduce the *compactified* gauge transformations [for example, see I. I. Kogan and A. Kovner, Phys. Rev. D **53**, 4510 (1996); **51**, 1948 (1995)] defined as

$$a_\mu\to a_\mu+\partial_\mu\theta_{\rm reg}+\partial_\mu\theta_{\rm sing},$$

where $\theta_{\rm reg}(x)$ describes the ordinary smooth or regular gauge transformations. These familiar gauge transformations are extended to include singular gauge transformations described by $\theta_{\rm sing}(x)$, itself determined from

$$(\partial\times\partial\theta_{\rm sing}(x))_\mu=2\pi\sum_{k=1}^{\mathcal{N}}\int_L dx_{k\mu}\delta(x-x_k),$$

where the sum runs over an arbitrary number $\mathcal{N}$ of unit-flux Dirac strings and $\int_L$ denotes a line integral. We now inquire whether the theory (6) is invariant under such compactified gauge transformations when applied to a_μ. The answer is "yes." However, when the same type of transformation is separately applied to v_μ, the answer is "no," due to its Meissner coupling to BdG fermions. Hence the labels. Note that both $v_{A\mu}$ and $v_{B\mu}$ are noncompact in the same sense.

[28] For computational convenience we define $\mathbf{V}_s$ as twice the conventional superfluid density; see Ref. 9.

[29] An exception to this general situation is the Anderson gauge $\varphi_A=\varphi$, $\varphi_B=0$ or $\varphi_A=0$, $\varphi_B=\varphi$ [P. W. Anderson, cond-mat/9812063 (unpublished)], which has least desirable properties in this regard, resulting in $\mathbf{v}=\mathbf{a}$ $(v_\mu=a_\mu)$ or an *infinitely* strong coupling between $\mathbf{v}$ and $\mathbf{a}$. It is then impossible to decouple $\mathbf{v}$ (v_μ) and $\mathbf{a}$ (a_μ), even in the RG sense. Consequently, any attempt to construct the effective theory by using the Anderson gauge is unlikely to be successful. To illustrate this, note that in the Anderson gauge (also in a general "$A-B$" gauge) the system has *finite* overall magnetization: $M=n_\uparrow-n_\downarrow=\langle\bar\psi_\uparrow\psi_\uparrow\rangle-\langle\bar\psi_\downarrow\psi_\downarrow\rangle=\langle\tilde\psi^\dagger_\uparrow\tilde\psi_\uparrow\rangle-\langle\tilde\psi^\dagger_\downarrow\tilde\psi_\downarrow\rangle\neq0$. In contrast, in the FT gauge (4) and (11), $M=0$, as it must be. Again, this preference for the FT gauge is a consequence of the intricate nature of the "coarse-graining" process—if the gauge fields are defined on *discrete* vortex configurations $\{\mathbf{r}_i(\tau)\}$, Eqs. (7) and (8), the magnetization vanishes with any choice of gauge.

[30] For the aficionados of gauge theories, this situation is reminiscent of the Landau gauge $(\partial\cdot A=0)$ being the fixed point among all covariant gauges of the Higgs-Abelian theory.

[31] At the risk of boring the reader, we reiterate that all the different singular gauges lead to precisely the same results if the partition function is evaluated *exactly*, by integrating over the *discrete* vortex variables $\{\mathbf{r}_i(\tau)\}$—it is only when we seek to perform this integration approximately, in terms of *distributed* quantities like $\rho(\mathbf{r})$ or $j_\mu(x)$, that the trouble arises.

[32] T. Senthil and M. P. A. Fisher, Phys. Rev. B **62**, 7850 (2000).

[33] D. A. Bonn *et al.*, Nature (London) **414**, 887 (2001).

[34] Z. Tešanović, Phys. Rev. B **59**, 6449 (1999); see Appendices A and B.

[35] A. Nguyen and A. Sudbø, Phys. Rev. B **60**, 15 307 (1999).

[36] Irrelevant, but *dangerously* so, since it breaks the discrete symmetry of the fermion Lagrangian.

[37] S. H. Simon and P. A. Lee, Phys. Rev. Lett. **78**, 1548 (1997).

[38] V. Gusynin, A. Hams, and M. Reenders, Phys. Rev. D **63**, 045025 (2001).

[39] See, e.g., M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory* (Addison-Wesley, Reading, MA, 1995).

[40] M. Reenders, cond-mat/0110168 (unpublished).

[41] D. H. Kim, P. A. Lee, and X.-G. Wen, Phys. Rev. Lett. **79**, 2109 (1997); D. H. Kim and P. A. Lee, Ann. Phys. (N.Y.) **272**, 130 (1999).

[42] T. W. Appelquist and U. Heinz, Phys. Rev. D **24**, 2169 (1981); D. Atkinson, P. W. Johnson, and P. Maris, *ibid.* **42**, 602 (1990).

[43] M. R. Pennington and D. Walsh, Phys. Lett. B **253**, 246 (1991).

[44] L. S. Brown, *Quantum Field Theory* (Cambridge University Press, Cambridge, 1992), pp. 521–524.

[45] D. G. Boulware, Phys. Rev. **151**, 1024 (1966).

[46] We note that, superficially, Eq. (78) resembles a result one would obtain by decoupling the fermion fields from the line integral in

Eq. (76) and then evaluated the latter in the leading-order cummulant expansion. However, it is easy to see that such a procedure would yield a *wrong sign* for $F(r)$ and the resulting expression would not be gauge invariant. A more detailed consideration based on a careful treatment of gauge invariance (Ref. 44) leads to correct result, Eq. (78).

[47] It is not immediately obvious what such regularization might be within QED$_3$ itself—it might involve replacing the straight line integral with a more ramified path or a family of paths. Within our theory, however, it is clear that such a regularization ultimately reduces Eq. (76) to the original expression (73), with discrete vortex variables.

[48] This result was originally obtained by V.P. Gusynin (unpublished). We are indebted to him for generously sharing it with us and especially for drawing our attention to Ref. 44.

[49] T. Valla *et al.*, Phys. Rev. Lett. **85**, 828 (2000).

[50] O. K. Andersen, O. Jepsen, A. I. Liechteinstein, and I. I. Mazin, Phys. Rev. B **49**, 4145 (1994).

[51] W. Rantner and X.-G. Wen, Phys. Rev. Lett. **86**, 3871 (2001).

[52] W. Rantner and X.-G. Wen, cond-mat/0010378 (unpublished).

[53] D. V. Khveshchenko, cond-mat/0112202 (unpublished).

[54] References 51 and 52 quote a positive value for η'. We find that this is a result of an overall sign error and η' should be negative, in agreement with Ref. 53. We are indebted to W. Rantner and D. V. Khveshchenko for communications regarding this issue and for their many insightful comments concerning gauge invariance.

[55] G. Leibbrandt and K. A. Richardson, Phys. Rev. D **46**, 2578 (1992).

[56] L. Baulieu and D. Zwanziger, Nucl. Phys. B **548**, 527 (1999).

[57] A. Melikyan and Z. Tešanović (unpublished).

[58] I. F. Herbut and Z. Tešanović, Phys. Rev. Lett. **76**, 4588 (1996).

[59] See, e.g., M. R. Norman, A. Kaminski, J. Mesot, and J. C. Campuzano, Phys. Rev. B **63**, 140508 (2001).

[60] M. Randeria, in *Superconductivity*, Proceedings of the International School of Physics ‘‘Enrico Fermi,’’ Course CXXXVI, edited by G. Iadonisi, J. R. Schrieffer, and M. L. Chiafalo (IOS Press, 1998), pp. 53–75, and references therein. See also N. Trivedi and M. Randeria, Phys. Rev. Lett. **75**, 312 (1995).

[61] V. M. Loktev, R. M. Quick, and S. G. Sharapov, Phys. Rep. **349**, 1 (2001).

[62] Q. Chen, K. Levin, and I. Kosztin, Phys. Rev. B **63**, 184519 (2001); and references therein.

[63] C. Berthod and B. Giovannini, Phys. Rev. Lett. **87**, 277002 (2001).

[64] S. C. Zhang, Science **275**, 1089 (1997).

[65] S. Chakravarty, R. B. Laughlin, D. K. Morr, and C. Nayak, Phys. Rev. B **63**, 094503 (2001).

[66] M. Vojta, Y. Zhang, and S. Sachdev, Phys. Rev. B **62**, 6721 (2000).

[67] QED$_3$ is a *superrenormalizable* field theory; i.e., it does not suffer from ultraviolet divergences. See P. Maris, Phys. Rev. D **52**, 6087 (1995).

[68] H. Kleinert, *Gauge Fields in Condensed Matter* (World Scientific, Singapore, 1989), Vol. I.